

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 2002, 11:47:52 ; Search time 25 Seconds

(without alignments)
248,858 Million cell updates/sec

Title: US-09-619-380-2

Perfect score: 798

Sequence: 1 MGVPNPNFNSWVLPGRLAGLA.....GSIETFEDEKAVFQFQRTK 150

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	137.5	17.2	551 1	CC14_YEAST
2	119.5	15.0	989 1	PPP3_DICCI
3	115.5	14.5	663 1	DUS8_MOUSE
4	115	14.4	597 1	MCE1_HUMAN
5	115	14.4	597 1	MCE1_MOUSE
6	114.5	14.3	625 1	DUS8_HUMAN
7	114	14.3	168 1	PNP_NPVAC
8	114	14.3	665 1	DUSX_HUMAN
9	113	14.2	573 1	MCE1_CAEL
10	111.5	14.0	340 1	DUSC_HUMAN
11	107	13.4	220 1	PNP1_NPVOP
12	107	13.4	384 1	DUS9_HUMAN
13	104	13.0	276 1	PNP3_CHLEU
14	104	13.0	595 1	PNP6_HUMAN
15	104	13.0	595 1	PNP6_MOUSE
16	103.5	13.0	159 1	Y215_METYA
17	99	12.4	360 1	PNP7_HUMAN
18	98.5	12.3	699 1	PNPE_MOUSE
19	98	12.3	700 1	PNPE_HUMAN
20	96.5	12.1	147 1	Y220_METYA
21	95.5	12.0	359 1	PNP7_RAT
22	94.5	11.8	280 1	DUS7_RAT
23	94.5	11.8	320 1	DUS7_HUMAN
24	94.5	11.8	320 1	DUS7_MOUSE
25	94.5	11.8	377 1	PNP2_DICDI
26	94	11.8	381 1	DUS6_HUMAN
27	94	11.8	381 1	DUS6_MOUSE
28	94	11.8	381 1	DUS6_RAT
29	94	11.8	802 1	PNP8_MOUSE
30	93	11.7	458 1	PNP1_HUMAN
31	92.5	11.6	1705 1	PNP0_MOUSE
32	92	11.5	369 1	PNP5_RAT
33	92	11.5	432 1	PNP1_RAT

34	92	11.5	434 1	PTN1_CHICK	013016 gallus galli
35	91.5	11.5	802 1	PTN1_HUMAN	P18433 homo sapien
36	91.5	11.5	829 1	PTN1_MOUSE	P18052 mus musculu
37	91.5	11.5	845 1	CSW_DROME	P29349 drosophila
38	91.5	11.5	1711 1	PTN1_RAT	O64612 ratu norv
39	91	11.4	913 1	PTN3_HUMAN	P26045 homo sapien
40	90.5	11.3	415 1	PTN2_HUMAN	P17706 homo sapien
41	90.5	11.3	541 1	PTN5_MOUSE	P54830 mus musculu
42	90	11.3	432 1	PTN1_MOUSE	P35821 mus musculu
43	90	11.3	489 1	MSG5_YEAST	P38590 saccharomyc
44	89.5	11.2	363 1	PTN2_RAT	P35233 ratu norv
45	89	11.2	403 1	PTN1_HUMAN	O00633 homo sapien

ALIGNMENTS

RESULT 1
ID CC14_YEAST STANDARD: PRT: 551 AA.
AC 000684;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable protein-tyrosine phosphatase CDC14 (EC 3.1.3.48).
GN CDC14 OR YFR028C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92283835; PubMed=1597462;
RA Wan J., Xu H., Grunstein M.;
RT "CDC14 of Saccharomyces cerevisiae. Cloning, sequence analysis, and
transcription during the cell cycle.";
RL J. Biol. Chem. 267:11274-11280(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RA MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagihara H., Shibata T., Ozawa M.,
RA Sasagawa S.-I., Sasagawa M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
Saccharomyces cerevisiae.";
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RA MEDLINE=96287654; PubMed=8686381;
RA Eki T., Naitou M., Hagihara H., Ade M., Ozawa M., Sasagawa S.-I.,
RA Sasagawa M., Tsuchiya Y., Shibata T., Watanabe K., Ono A.,
RA Yamazaki M., Tashiro H., Hanaka F., Murakami Y.;
RT "Fifteen open reading frames in a 30.8 kb region of the right arm of
chromosome VI from Saccharomyces cerevisiae.";
RL Yeast 12:177-190(1996).
RN [4]
RP FUNCTION: PROBABLE PROTEIN-TYROSINE PHOSPHATASE THAT MAY BE
INVOLVED IN CHROMOSOME SEGREGATION.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O -> protein
tyrosine + phosphate.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
TYROSINE PHOSPHATASE FAMILY. CDC14 SUBFAMILY.
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN EXTENSIVELY IN
THE N-TERMINAL DUE TO A FRAMESHIFT.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

```

CC EMBL: M61194; AAA34477.1; ALT_FRAME.
DR EMBL: D50617; BAA09267.1; -.
DR PIR: A42784; A42784.
DR SGD: S0001924; CDC14.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC.1.
DR SMART: SM00195; DSPC.1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR Cell cycle; Hydrolase.
FT ACT_SITE 283 283
SQ SEQUENCE 551 AA; 61906 MW; 4EB3985DFA3FD823 CRC64;

Query Match 17.28; Score 137.5; DB 1; Length 551;
Best Local Similarity 28.0%; Pred. No. 2,9e-06;
Matches 30; Conservative 20; Mismatches 34; Indels 23; Gaps 3;

QY 27 HYGFLLDGLVRLHSLTERGPPHSDCPGLTLRLRLPDCPPAPDQIDRFVQIVDEANNA 86
DB 238 HFE---DIGQHLDLIEDG-----TCDDLSI-----VKNFVGAETTIK 274

QY 87 RGEAVGVHCAIGFGRTGTMACYLVKERGLAAGDAIAEIRLRPGSI 133
DB 275 RGGKIAVHCAGIGRTGCLIGAHLYTFTANECIGELFIRGMV 321

RESULT 2
PTP3_DICDI STANDARD; PRT; 989 AA.
ID PTP3_DICDI
AC P54637;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Protein-tyrosine phosphatase 3 (EC 3.1.3.48) (Protein-tyrosine-
DE phosphate phosphohydrolase 3).
GN (PTP3 OR PTP3) AND (PTP3 OR PTP3).
OS Dictyostellum discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AX3;
RX MEDLINE=96189126; PubMed=8628311;
RA Gamber M., Howard P.K., Hunter T., Firtel R.A.;
RT "Multiple roles of the novel protein tyrosine phosphatase PTP3 during
RT Dictyostellum growth and development."
RL Mol. Cell. Biol. 16:2431-2444(1996).
CC -1- FUNCTION: SEEMS TO DEPHOSPHORYLATE A PROTEIN OF 130 kDa (P130).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: IN THE ANTERIOR-LIKE AND PRESTALK CELL TYPES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT MODERATE LEVELS DURING GROWTH
CC AND DEVELOPMENT.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U38197; AAC47041.1; -.
DR HSP: 006124; ZSNP.
DR DictyDb: DD0111; PTPC1.
DR DictyDb: DD0777; PTPC2.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; Tyr_PP.

```

```

DR Pfam: PF00102; Y-phosphatase; 1.
DR PRINTS: PR00700; PTPPHPTASE.
DR SMART: SM00194; PTPC.1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 1.
DR Hydrolase.
FT ACT_SITE 649 649
SQ SEQUENCE 989 AA; 109995 MW; 9371105AF80974AF CRC64;

Query Match 15.03; Score 119.5; DB 1; Length 989;
Best Local Similarity 28.0%; Pred. No. 0.00036;
Matches 37; Conservative 22; Mismatches 48; Indels 25; Gaps 5;

QY 33 DLGVR-HLVSLTERGPPHSDCPGLTLRLRLPDCPPAPDQIDRFVQIVDEANNA 89
DB 583 DIGVSLHLVKKGEFFREV-----VLLHTQWDCCAPRSSIIRLSVWNTFKRGS 637

QY 90 A-----GVHCAIGFGRTGTMACYL-----VKEGLAAGDAIAEIRLRPGSI 132
DB 638 AKNTNGFVIVHCSGIGRTGCLIGAHLYTFTANECIGELFIRGMV 697

QY 133 IETVEOKAVFO 144
DB 698 VQIDQYTFIFK 709

RESULT 3
DUS8_MOUSE STANDARD; PRT; 663 AA.
ID DUS8_MOUSE
AC O09112;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Neuronal tyrosine threonine phosphatase 1).
GN DUSP8 OR NTP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE=96311565; PubMed=8733137;
RA Theodosiou A.M., Rodrigues N.R., Nesbitt M.A., Ambrose H.J.,
RA Peterson H., McEldan-Arnold E., Boyd Y., Leversha M.A., Owen N.,
RA Blake D.J., Ashworth A., Davies K.E.;
RT "A member of the MAP kinase phosphatase gene family in mouse
RT containing a complex trinucleotide repeat in the coding region."
RL Hum. Mol. Genet. 5:675-684(1996).
CC -1- FUNCTION: THIS PROTEIN SHOWS BOTH ACTIVITY TOWARD TYROSINE-PROTEIN
CC PHOSPHATE AS WELL AS WITH SERINE/THREONINE-PROTEIN PHOSPHATE (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN BRAIN AND LUNG.

```


FT	ACT_SITE	126	126	RNA 5'-TRIPHOSPHATASE (BY SIMILARITY).
FT	ACT_SITE	294	294	GUANYLATION SITE.
FT	VARSPIC	424	446	MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT	VARSPIC	481	597	MISSING (IN ISOFORM 3).
FT	VARSPIC	504	597	TREKQYDNKLTIECFENNNSVFWFRDRKSPNAYNTAMA
FT				VCSNSINVTETEMLEFFEDRCFASOGGRKHHPDELM
FT				PPPPKRRPLT -> CLFIRSYELDVLTLSIHQNNNN
FT				OHITSSSTGG (IN ISOFORM 4).
FT	MUTAGEN	294	294	K->: LOSS OF GTPASE ACTIVITY.
FT	MUTAGEN	299	299	R->: LOSS OF GTPASE ACTIVITY.
FT	MUTAGEN	345	345	E->: LOSS OF GTPASE ACTIVITY.
FT	MUTAGEN	458	458	K->: LOSS OF GTPASE ACTIVITY.
FT	MUTAGEN	460	460	K->: LOSS OF GTPASE ACTIVITY.
FT	CONFLICT	30	30	M -> I (IN REF. 1).
FT	CONFLICT	484	484	Q -> P (IN REF. 1).
SQ	SEQUENCE	597 AA;	68556 MW;	SICEECIBI90603DE CRC64;
Query Match				
Best Local Similarity		14.4%;	Score 115;	DB 1; Length 597;
Matches 24; Conservative		30.0%;	Pred. No. 0.00058;	
		18;	Mismatches 36;	Indels 2; Gaps 1
QY	69	PAPQIDREFQVIDEANARG--EAGVCHALGFGFTGMACYLKERGLAAGDAIEIR	126	
DB	98	PTTNTETFTLIERFERNPELIGVCTGFNGFTGLICAFLEKMDMSIAAVATPA	157	
QY	127	RLRGSITETYEQEKARPOFY	146	
DB	158	QARPPGTYKGDYKLELFERY	177	
RESULT 5				
MCEL_MOUSE				
ID	MCEL_MOUSE	STANDARD;	PRT;	597 AA.
AC	055236;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	mRNA capping enzyme (MCE) [Includes: polynucleotide 5'-			
DE	triphosphatase (EC 3.1.3.33) (mRNA 5'-triphosphatase) (trase); mRNA			
DE	guanylyltransferase (EC 2.7.7.50) (GTP--RNA guanylyltransferase)			
DE	(trase)]			
GN	RNGT OR CAPA.			
OS	Mus musculus (Mouse).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euteleota; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OC	NCBI_TaxID=10090;			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-294.			
RA	MEDLINE=98058741; Pubmed=9371772;			
RA	Yue Z., Maldonado E., Piluluta R., Cho H., Reinberg D., Shatkin A.J.;			
RA	"Mammalian capping enzyme complements mutant Saccharomyces cerevisiae			
RT	lacking mRNA guanylyltransferase and selectively binds the elongating			
RT	form of RNA polymerase II.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 94:12898-12903(1997).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=98070332; Pubmed=9407024;			
RA	McCricken S., Fong N., Rosonina E., Yankulov K., Brothers G.,			
RA	Siderovski D., Hessel A., Foster S., Shuman S., Bentley D.L.;			
RA	5'-capping enzymes are targeted to pre-mRNA by binding to the			
RT	phosphorylated carboxy-terminal domain of RNA polymerase II.";			
RT	Genes Dev. 11:3306-3318(1997).			
RL	[3]			
RP	CHARACTERIZATION, AND MUTAGENESIS.			
RA	MEDLINE=98445353; Pubmed=9770468;			
RA	Men Y., Yue Z., Shatkin A.J.;			
RT	"Mammalian capping enzyme binds RNA and uses protein tyrosine			
RT	phosphatase mechanism.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 95:12226-12231(1998).			
CC	-1- FUNCTION: BIFUNCTIONAL mRNA CAPPING ENZYME EXHIBITING RNA 5'-			
CC	TRIPHOSPHATASE ACTIVITY IN THE N-TERMINAL PART AND RNA			
CC	GUANYLYLTRANSFERASE ACTIVITY IN THE C-TERMINAL PART. CATALYZES THE			
CC	FIRST TWO STEPS OF CAP FORMATION: BY REMOVING THE GAMMA-PHOSPHATE			

	CC	FROM THE 5'-TRIPHOSPHATE END OF NASCENT RNA TO YIELD A
	CC	DIPHOSPHATE END, AND BY TRANSFERRING THE GMP MOIETY OF GTP TO THE
	CC	5-DIPHOSPHATE TERMINUS.
	CC	-1 CATALYTIC ACTIVITY: 5'-phosphopolynucleotide + H(2)O =
	CC	polynucleotide + phosphate.
	CC	-1 CATALYTIC ACTIVITY: GTP + (5')pp-pur-mRNA = diphosphate +
	CC	G(5'),PP-pur-mRNA.
	CC	-1 ENZYME REGULATION: RNA TRIPHOSPHATASE ACTIVITY IS INHIBITED BY
	CC	VANADATE, IODOACETATE, AND MAGNESIUM.
	CC	-1 SUBCELLULAR LOCATION: Nuclear.
	CC	-1 MISCELLANEOUS: BINDS TO THE ELONGATING PHOSPHORYLATED FORM OF RNA
	CC	POLYMERASE II. CAN INTERACT INDIRECTLY BY BINDING TO POL. II C-
	CC	TERMINAL DOMAIN AND DIRECTLY BY RNA BINDING. THE GPASE DOMAIN,
	CC	RATHER THAN THE TPASE DOMAIN MEDIATES THESE INTERACTIONS.
	CC	-1 SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE NON-RECEPTOR
	CC	CLASS OF THE PROTEIN-TYROSINE PHOSPHATASE FAMILY.
	CC	-1 SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE EUKARYOTIC
	CC	GPASE FAMILY.

	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
	CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
	CC	The European Bioinformatics Institute. There are no restrictions on its con-
	CC	tent by non-profit institutions as long as its content is in no way
	CC	modified and this statement is not removed. Usage by and for commercial
	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
	CC	or send an email to license@isb-sib.ch).
	CC	-----
DR	EMBL;	AFO25653; AA891558.1; -
DR	EMBL;	AFO34568; AA888903.1; -
DR	MGI:	MGI:1328041; Ragtl.
DR	InterPro;	IPR000340;.DS.phosphatase.
DR	InterPro;	IPR000387; TYR_phosphatase.
DR	InterPro;	IPR001339; mRNA_cap_enzyme.
DR	Pfam;	PF01331; mRNA_cap_enzyme.1.
DR	SMART;	SMO0012; PTpc_DSPC.1.
DR	PROSITE;	PS00383; TYR_PHOSPHATASE_1; 1.
DR	PROSITE;	PS00566; TYR_PHOSPHATASE_2; 1.
KM	mRNA processing;	mRNA capping; Transferase; Nucleotidyltransferase;
KW	Hydrolase;	Multi-functional enzyme; Nuclear protein.
FT	DOMAIN	1 212
FT	DOMAIN	229 597
FT	DOMAIN	195 205
FT	ACT_SITE	126 126
FT	ACT_SITE	294 294
FT	ACT_SITE	36 36
FT	MUTAGEN	66 66
FT	MUTAGEN	110 110
FT	MUTAGEN	125 125
FT	MUTAGEN	126 126
FT	MUTAGEN	132 132
FT	MUTAGEN	133 133
FT	MUTAGEN	138 138
FT	MUTAGEN	168 168
FT	MUTAGEN	290 290
FT	MUTAGEN	294 294
FT	MUTAGEN	315 315
FT	MUTAGEN	315 315
FT	MUTAGEN	530 530
FT	MUTAGEN	530 530
FT	MUTAGEN	533 533
FT	MUTAGEN	533 533
FT	MUTAGEN	537 537
FT	MUTAGEN	537 537
SEQ	SEQUENCE	597 AA; 68684 MW; BA9SE111889DEA7 CRC64;

Query Match 14.4%; Score 115; DB 1; Length 597;
 Best Local Similarity 30.0%; Pred. No. 0.00058;
 Matches 24; Conservative 18; Mismatches 36; Indels 2; Gaps 1;

OY 69 PAPPDIDRVOQVDEANARG--EAVGHCALGFGRTGTMACYLKERGLAAGDAIAEIR 126
 Db 98 PPTMETETIRLCERENSPPELLIGHCTGCFNRTGFLICAPLVEKMDWSTEAATVTA 157
 OY 127 RLRPSTIEYEDKEANFOFY 146
 Db 158 QARPEITYGDIYKLELERRY 177

RESULT 6

DUS8_HUMAN STANDARD; PRT; 625 AA.

AC Q1302; 15-JUL-1999 (rel. 38, Created)
 DT 15-JUL-1999 (rel. 38, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
 DE (Dual specificity protein phosphatase hnh-5).
 GN DUSP8 OR VHS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=96009533; PubMed=7561881;
 RA Martell K.J., Seasholtz A.F., Kwak S.P., Clemens K.K., Dixon J.E.;
 RT "hnh-5: a protein tyrosine phosphatase abundant in brain that
 RT inactivates mitogen-activated protein kinase.";
 RL J. Neurochem. 65:1823-1833(1995).
 CC -1- FUNCTION: THIS PROTEIN SHOWS BOTH ACTIVITY TOWARD TYROSINE-PROTEIN
 CC PHOSPHATE AS WELL AS WITH SERINE/THREONINE-PROTEIN PHOSPHATE (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- TISSUE SPECIFICITY: ABUNDANT IN BRAIN, HEART, AND SKELETAL MUSCLE.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 RHODANESE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U27193; AAA83151.1; -
 DR HSSP: Q16828; IMKP.
 DR Genew: HGNC:3074; DUSP8.
 DR MIM: 602038; -
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR001763; Rhodanease-like.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSPC; 1.
 DR SMART: SM00195; DSPC; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS00566; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS00564; TYR_PHOSPHATASE_DUAL; 1.
 KM Hydrolyase; Nuclear protein.
 FT DOMAIN 13 136
 FT DOMAIN 162 430
 CC DOMAIN 310 550 PROTEIN-TYROSINE PHOSPHATASE.
 CC PRO-RICH.

FT ACT_SITE 246 246 BY SIMILARITY.
 SQ SEQUENCE 625 AA; 65840 MW; DCBEAL4487219666 CRC64;

Query Match 14.3%; Score 114.5; DB 1; Length 625;
 Best Local Similarity 23.5%; Pred. No. 0.00069;
 Matches 40; Conservative 27; Mismatches 64; Indels 39; Gaps 6;

OY 8 FSNVPLGRLAG-----LALRLPA-----HYQFLDGVNHLVSLTERGP 48
 Db 129 FSSCFPGLCBGRPALLPMSLSQCLPVPVSGITRLPILHYLGSDQKDVNKLDMTONGIS 188
 OY 49 H----SDSCPG-----LTLRLRIP--DFCPAPDIDRFVOVDENARGAAGVHCA 97
 Db 189 IYVNASNSCPKPDFICESRFMRVPINDNCEKILPMLDLSIEFTIDAKLSSCOVIVHCLA 248
 OY 98 GFGRTGTMLACYLKERGLAAGDAIAEIRLRP-----GSIETEQ 138
 Db 249 GISRSATVIAIAYIMKTQMSDDAYRFVADRRPSISPNFNIGQLLEYER 298

RESULT 7

P1P_NPVAC STANDARD; PRT; 168 AA.

AC P24656; 01-MAR-1992 (rel. 21, Created)
 DT 01-NOV-1995 (rel. 32, Last sequence update)
 DT 01-NOV-1995 (rel. 32, Last annotation update)
 DE Protein tyrosine phosphatase (EC 3.1.3.48) (BVP).
 GN PTP.
 OS Autographa californica nuclear polyhedrosis virus (AcNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_Taxid=46015;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91132123; PubMed=1993873;
 RA Tilarakate N., Hardin S.E., Weaver R.F.;
 RT "Nucleotide sequence and transcript mapping of the HindIII F region
 RT of the Autographa californica nuclear polyhedrosis virus genome.";
 RL J. Gen. Virol. 72:285-291(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C6;
 RX MEDLINE=92024079; PubMed=1926775;
 RA Possee R.D., Sun T.P., Howard S.C., Ayres M.D., Hill-Berkins M.,
 RA Gearing K.L.;
 RT "Nucleotide sequence of the Autographa californica nuclear
 RT polyhedrosis 9.4 kbp EcoRI-I and -R (polyhedrin gene) region.";
 RL Virology 185:229-241(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C6;
 RX MEDLINE=94303173; PubMed=8030224;
 RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
 RT "The complete DNA sequence of Autographa californica nuclear
 RT polyhedrosis virus.";
 RL Virology 202:586-605(1994).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=93186773; PubMed=8444848;
 RA Sheng Z., Charbonneau H.;
 RT "The baculovirus Autographa californica encodes a protein tyrosine
 RT phosphatase.";
 RL J. Biol. Chem. 268:4728-4733(1993).
 CC -1- FUNCTION: ALSO DEPHOSPHORYLATES SERYL AND THREONYL RESIDUES.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- MISCELLANEOUS: PROBABLY EXPRESSED LATE IN INFECTION.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. CDC14 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION, AND MUTAGENESIS OF CYS-124.
RX MEDLINE-97344078; Pubmed-9200605;
RA Takagi T., Moore C.R., Diehn F., Buratowski S.;
RT "An RNA 5'-triphosphatase related to the protein tyrosine
RT phosphatases";
RT Cell 89:867-873(1997).
CC -1- FUNCTION: BIFUNCTIONAL RNA CAPPING ENZYME EXHIBITING RNA 5'-
CC TRIPHOSPHATASE ACTIVITY IN THE N-TERMINAL PART AND RNA
CC GUANYLYLTRANSFERASE ACTIVITY IN THE C-TERMINAL PART. CATALYZES THE
CC FIRST TWO STEPS OF CAP FORMATION: BY REMOVING THE GAMMA-PPHOSPHATE
CC FROM THE 5'-TRIPHOSPHATE END OF NASCENT RNA TO YIELD A
CC DIPHOSPHATE END, AND BY TRANSFERRING THE GMP MOIETY OF GTP TO THE
CC 5'-DIPHOSPHATE TERMINUS.
CC -1- CATALYTIC ACTIVITY: 5'-phosphopolynucleotide + H(2)O =
CC polynucleotide + phosphate.
CC -1- CATALYTIC ACTIVITY: GTP + (5')pp-pur-mRNA = diphosphate +
CC G(5')pp-pur-mRNA.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- INDUCTION: INHIBITED BY MAGNESIUM.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE NON-RECEPTOR
CC CLASS OF THE PROTEIN-TYROSINE PHOSPHATASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE EUKARYOTIC
CC GTPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF003925; AB61344.1; -
DR EMBL: Z75525; CA99765.1; ALT_INIT.
DR Wormpep: C03D6.3; CE15578.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR001339; mRNA_cap_enzyme.
DR Pfam: PF01331; mRNA_cap_enzyme; 1.
DR SMART: SM00012; PTPC_DsPc; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR mRNA processing; mRNA capping; Transference; Nucleotidyltransferase;
DR Hydrolyase; Multifunctional enzyme; Nuclear protein.
FT DOMAIN 1 212 TRASE.
FT FT 229 573
FT ACT_SITE 124 124 RNA 5'-TRIPHOSPHATASE.
FT ACT_SITE 299 299 GUANYLYLATION SITE (BY SIMILARITY).
FT MUTAGEN 124 124 C->S: LOSS OF ACTIVITY.
FT MUTAGEN 124 124 C->A: LOSS OF ACTIVITY.
SQ SEQUENCE 573 AA; 66336 MW; F50ACAS9F1815F47 CRC64;

Query Match 14.2%; Score 113; DB 1; Length 573;
Best Local Similarity 36.0%; Pred. No. 0.00088;
Matches 31; Conservative 12; Mismatches 37; Indels 6; Gaps 3;

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dual specificity protein phosphatase 12 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Dual specificity tyrosine phosphatase YWH1).
GN DUSP12.
OS "Homo sapiens (Human)".
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID-9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE-99377030; Pubmed-10446167;
RA Wada M., Manning E.R., Orth K., Dixon J.E.;
RT Identification of the human YWH1 protein-tyrosine phosphatase
RT orthologue reveals a novel zinc binding domain essential for in vivo
RT function";
RT J. Biol. Chem. 274:23991-23995(1999).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- COFACTOR: BINDS 2 MOLES OF ZINC PER MOLE OF PROTEIN.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: UBIQUITOUS, HIGHEST EXPRESSION IN SPLEEN,
CC TESTIS, OVARY, AND PERIPHERAL BLOOD LEUKOCYTES AND LOWER
CC EXPRESSION IN LIVER AND LUNG.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF19226; AAD51134.1; -
DR HSSP: 016828; 1MRP.
DR GeneW: HGNC:3067; DUSP12.
DR MIM: 604835; -
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF00782; DSPC; 1.
DR SMART: SM00195; DSPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; FALSE NEG.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; UNKNOWN_1.
DR Hydrolyase; zinc; Metal-binding.
FT DOMAIN 88 160 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 115 115 BY SIMILARITY.
SQ SEQUENCE 340 AA; 37687 MW; 56B52192B42C73EB CRC64;

Query Match 14.0%; Score 111.5; DB 1; Length 340;
Best Local Similarity 28.4%; Pred. No. 0.00069;
Matches 40; Conservative 26; Mismatches 64; Indels 11; Gaps 6;

ID PTP1_NPVPD STANDARD; PRT; 220 AA.
AC 010274;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Putative protein-tyrosine phosphatase 1 (EC 3.1.3.48).
OS Orygia pseudotsugata multicaudal polyhedrosis virus (OPMVV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
NC Nucleopolyhedrovirus.
RX NCBI_TaxId=164623;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9271300; PubMed=9126251;
RA Atkins C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
RA Rohmann G.F.;
RT "The sequence of the Orygia pseudotsugata multicaudal nuclear
RT polyhedrosis virus genome.";
RL Virology 229:381-399(1997).
CC -1- FUNCTION: COULD BE INACTIVE AS THE ACTIVE SITE CYSTEINE IS
CC MODIFIED TO TRYPTOPHAN.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. CDC14 SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U75930; AAC5909.1; -
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR SMART; SM00012; PTPc_DSPc.1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
KW Hydrolase.
KW HYDROLASE.
FT SITE 162 EQUIVALENT OF ACTIVE SITE CYS.
SQ SEQUENCE 220 AA; 25234 MW; D3FC093F195D445 CRC64;
Query Match 13.4%; Score 107; DB 1; Length 220;
Best Local Similarity 32.9%; Pred. No. 0.0012;
Matches 27; Conservative 13; Mismatches 40; Indels 2; Gaps 1;
QY 55 GLTHRLRIPDPCPPAPDOIDRFVQYVDENARGEA--VGYHCAIGRGRTGTMACYLVK 112
DB 120 GLTKKIRVPGRAVDDDIYAETVDEFFRRCPTMLVAVMTHTGLNRSGLVCRWAVE 179
QY 113 ENGLAAGDAIAEIRRLRPGSIE 134
DB 180 RLGVSPDAIARFETARGHKIE 201
RESULT 12
DUS9_HUMAN STANDARD; PRT; 384 AA.
ID DUS9_HUMAN
AC 099956;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Dual specificity protein phosphatase 9 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 4) (MAP kinase
DE phosphatase 4) (MKP-4).
GN DUSP9 OR MKP4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=97184169; PubMed=9030581;
RA Mada M., Boshett U., Smith A., Antonsen B., Gillieron C.,
RA Chabert C., Camps M., Martinou I., Ashworth A., Arkinstall S.;
RT "Molecular cloning and functional characterization of a novel
RT mitogen-activated protein kinase phosphatase, MKP-4.";
RL J. Biol. Chem. 272:5141-5151(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Platzer M.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INACTIVATES MAP KINASES. HAS A SPECIFICITY FOR THE ERK
CC FAMILY.
CC -1- CATALYTIC ACTIVITY: protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC
CC -1- SIMILARITY: CONTAINS 1 RHODANESE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Y08302; CA69610.1; -
DR EMBL; U52111; AAF74507.1; -
DR HSPB; Q16828; IMKP.
DR Genew; HGNC:3076; DUSP9.
DR MIM; 300134; -
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPc.1.
DR SMART; SM00195; DSPc.1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase.
KW DOMAIN 8 137 RHODANESE
FT DOMAIN 203 384 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 290 BY SIMILARITY.
SQ SEQUENCE 384 AA; 41867 MW; F8598CA95AB379B7 CRC64;
Query Match 13.4%; Score 107; DB 1; Length 384;
Best Local Similarity 25.5%; Pred. No. 0.0022;
Matches 36; Conservative 25; Mismatches 56; Indels 24; Gaps 4;
QY 2 GVOQPPNFS-----WYLPGRLAGIALPRLPAHYQFLDLSGRVLVSLTERGPHSDSC 54
DB 191 GATPPVGLRASFPVQILPNLYLSA--RDSANESLSAKIGIRIILVATPLPFFEFKNG 248
QY 55 GLTHRLRIPD-----PCPPAPDOIDRFVQYVDENARGEAVGAYGAGFGRTGTM 106
DB 249 DFHYKQIPDISDHSQNSLRFEPPA-----IEFIDELSONCGVLVHCLAGVSRSTVT 301
QY 107 ACYLVKRGSLAAGDAIAEIR 127
DB 302 VAYLMQKHLISLNDAYDLVKR 322
RESULT 13
PTP3_CHLEU
ID PTP3_CHLEU STANDARD; PRT; 276 AA.
AC 039491;
DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Dual specificity protein phosphatase (EC 3.1.3.48) (EC 3.1.3.16).
 GN VH-PTP13.
 OS Chlamydomonas eugametos.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadales; Chlamydomonas.
 OX NCBI_Taxid=3053;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-DTEX 10;
 RX MEDLINE=95323001; PubMed=7599654;
 RA Haring M.A., Siderius M., Jonak C., Hirt H., Walton K.M.,
 RA Musgrave A.;
 RT "Tyrosine phosphatase signalling in a lower plant: cell-cycle and
 RT oxidative stress-regulated expression of the Chlamydomonas eugametos
 RT VH-PTP13 gene."
 RL Plant J. 7:981-988(1995).
 CC -1- FUNCTION: COULD BE INVOLVED IN TYROSINE PHOSPHATASE SIGNALLING
 CC PATHWAYS, HAVING MAP-KINASES AS SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- DEVELOPMENTAL STAGE: NONDIVIDING GAMETES DID NOT EXPRESS THE VH-
 CC PTP13 GENE WHEREAS SYNCHRONOUSLY DIVIDING VEGETATIVE CELLS ONLY
 CC EXPRESSED VH-PTP13 IN THE EARLY G1-PHASE OF THE CYCLE.
 CC -1- INDUCTION: BY OXIDATIVE STRESS.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; X77938; CAA54910.1; -
 CC DR HSSP; P51452; 1VHR.
 CC DR InterPro; IPR000340; DS_Phasphatase.
 CC DR InterPro; IPR000387; TYR_phosphatase.
 CC DR Pfam; PF00782; DSPC; 1.
 CC DR SMART; SM00195; DSPC; 1.
 CC DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 CC DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
 CC DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
 CC KM Hydrolase.
 CC FT DOMAIN 89 276 CATALYTIC.
 CC FT ACT_SITE 172 172
 CC FT MUTAGEN 172 172 C->S; INACTIVE.
 CC SQ SEQUENCE 276 AA; 30310 MW; 034EF63951E0381 CRC64;
 CC
 CC Query Match 13.0%; Score 104; DB 1; Length 276;
 CC Best Local Similarity 27.6%; Pred. No. 0.0031;
 CC Matches 35; Conservative 21; Mismatches 57; Indels 14; Gaps 4;
 CC
 CC Oy 9 SWVLGRLAGLALPRNHYQFLDLGVNHLVSLFR-GPPISD-----SCGELTHARR 62
 CC Db 88 SIVVEGLKL-LSSECVESSELTKLGVTNHLVQVGLKPSHGRPTVLSLPLIDEGOD 146
 CC Oy 63 IPDFCPAPDQIDRFVQVDEANRGENAVGVCALGFGRTGTMACYLVEKGLAAGDAI 122
 CC Db 147 IVALPSC-----FQFLQQAQASGGVCLVHCLAGISASAVYAIYLMWTGMPTEAR 199
 CC Oy 123 AEIRRLR 129
 CC Db 200 AMVRRAR 206

ID PTNG_HUMAN STANDARD; PRT; 595 AA.
 AC P29350;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein-tyrosine phosphatase, non-receptor type 6 (EC 3.1.3.48)
 DE (protein-tyrosine phosphatase 1C) (PTP-1C) (Hematopoietic cell
 DE protein-tyrosine phosphatase) (SH-PTP1) (Protein-tyrosine phosphatase
 DE SHP-1).
 GN PRP6 OR PTP1C OR HCP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN (1)
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RX MEDLINE=92123209; PubMed=1732748;
 RA Yl T., Cleveland J.L., Ihle J.N.;
 RT "Protein tyrosine phosphatase containing SH2 domains:
 RT characterization, preferential expression in hematopoietic cells, and
 RT localization to human chromosome 12p12-p13."
 RL Mol. Cell. Biol. 12:836-846(1992).
 RN (2)
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 RC TISSUE-Breast;
 RX MEDLINE=91343005; PubMed=1652101;
 RA Shen S.H., Bastien L., Posner B.I., Chretien P.;
 RT "A protein-tyrosine phosphatase with sequence similarity to the SH2
 RT domain of the protein-tyrosine kinases."
 RL Nature 352:736-739(1991).
 RN (3)
 RP REVISIONS.
 RA Shen S.H., Bastien L., Posner B.I., Chretien P.;
 RL Nature 353:868-868(1991).
 RN (4)
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RX MEDLINE=92141214; PubMed=1736296;
 RA Plutsky J., Neel B.G., Rosenberg R.D.;
 RT "Isolation of a src homology 2-containing tyrosine phosphatase."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:1123-1127(1992).
 RN (5)
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RX MEDLINE=95394454; PubMed=7665165;
 RA Banville D., Stocco R., Shen S.H.;
 RT "Human protein tyrosine phosphatase 1C (PTP6) gene structure:
 RT alternate promoter usage and exon skipping generate multiple
 RT transcripts."
 RL Genomics 27:165-173(1995).
 RN (6)
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RX MEDLINE=96303695; PubMed=8723724;
 RA Ansari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,
 RA Malley T., Gibbs R.A.;
 RT "A gene-rich cluster between the CD4 and triosephosphate isomerase
 RT genes at human chromosome 12p13."
 RL Genome Res. 6:314-326(1996).
 RN (7)
 RP PHOSPHORYLATION.
 RX MEDLINE=95300784; PubMed=7781604;
 RA Li R.Y., Gails F., Ragab A., Ragab-Thomas J.M.F., Chap H.;
 RT "Tyrosine phosphorylation of an SH2-containing protein tyrosine
 RT phosphatase is coupled to platelet thrombin receptor via a pertussis
 RT toxin-sensitive heterotrimeric G-protein."
 RL EMBO J. 14:2519-2526(1995).
 RN (8)
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 248-399.
 RX MEDLINE=98447672; PubMed=9774441;
 RA Yang J., Liang X., Niu T., Meng W., Zhao Z., Zhou G.W.;
 RT "Crystal structure of the catalytic domain of protein-tyrosine
 RT phosphatase SHP-1."
 RL J. Biol. Chem. 273:28199-28207(1998).
 CC -1- FUNCTION: PLAYS A KEY ROLE IN HEMATOPOIESIS. THIS PRPASE ACTIVITY
 CC MAY DIRECTLY LINK GROWTH FACTOR RECEPTORS AND OTHER SIGNALING

CC PROTEINS THROUGH PROTEIN-TYROSINE PHOSPHORYLATION. THE SH2 REGIONS
CC MAY INTERACT WITH OTHER CELLULAR COMPONENTS TO MODULATE ITS OWN
CC PHOSPHATASE ACTIVITY AGAINST INTERACTING SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
CC tyrosine + phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, A LONG FORM (SHOWN HERE)
CC AND A SHORT FORM, ARE PRODUCED BY ALTERNATIVE SPLICING.
CC SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN HEMATOPOIETIC
CC CELLS.
CC -1- PFM: PHOSPHORYLATED ON SERINE AND TYROSINE.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M74903; AAA35963.1; -
DR EMBL: X62053; CAA43982.1; -
DR EMBL: M77273; AAA36610.1; -
DR EMBL: U15528; AAA82880.1; -
DR EMBL: U15536; AAA82880.1; JOINED.
DR EMBL: U15535; AAA82880.1; JOINED.
DR EMBL: U15534; AAA82880.1; JOINED.
DR EMBL: U15533; AAA82880.1; JOINED.
DR EMBL: U15532; AAA82880.1; JOINED.
DR EMBL: U15531; AAA82880.1; JOINED.
DR EMBL: U15530; AAA82880.1; JOINED.
DR EMBL: U15529; AAA82880.1; JOINED.
DR EMBL: U15528; AAA82879.1; JOINED.
DR EMBL: U15537; AAA82879.1; JOINED.
DR EMBL: U15535; AAA82879.1; JOINED.
DR EMBL: U15534; AAA82879.1; JOINED.
DR EMBL: U15533; AAA82879.1; JOINED.
DR EMBL: U15532; AAA82879.1; JOINED.
DR EMBL: U15531; AAA82879.1; JOINED.
DR EMBL: U15530; AAA82879.1; JOINED.
DR EMBL: U15529; AAA82879.1; JOINED.
DR EMBL: U47924; AAB51322.1; -
DR EMBL: S20825; S20825.
DR PIR: B42031; B42031.
DR PDB: 1GWI; 22-AUG-99.
DR PDB: 1FPR; 07-MAR-01.
DR Genbank: HGNC:9658; PTPN6.
DR MIM: 176883; -
DR InterPro: IPR000980; SH2.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; TYR_PP.
DR Pfam: PF00017; SH2; 2.
DR Pfam: PF00102; Y_phosphatase; 1.
DR PRINTS: PRO0700; TRYPPHATASE.
DR PRINTS: PRO0401; SH2DOMAIN.
DR PRODOM: PD000093; SH2; 2.
DR SMART: SM00194; PTPc; 1.
DR SMART: SM00252; SH2; 2.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 1.
DR PROSITE: PS50001; SH2; 2.
KW Hydrolase; SH2 domain; Repeat; phosphorylation; Alternative splicing;
KW 3D-structure.
FT DOMAIN 4 100 SH2 1.
FT DOMAIN 110 213 SH2 2.
FT DOMAIN 269 514 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 453 453

FT VARSPLIC 1 39 MISSING (IN SHORT ISOFORM).
FT VARSPLIC 40 44 SLAYR -> MLSRG (IN SHORT ISOFORM).
FT CONFLICT 1 3 MVR -> MLSRG (IN REF. 2).
FT CONFLICT 6 6 H -> L (IN REF. 5).
FT CONFLICT 86 86 L -> V (IN REF. 4).
FT CONFLICT 146 146 V -> E (IN REF. 5).
SQ SEQUENCE 595 AA; 67561 MW; 4D7736C21D3542D2 CXC64;
Query Match 13.0%; Score 104; DB 1; Length 595;
Best Local Similarity 27.5%; Pred. No. 0.0074;
Matches 39; Conservative 17; Mismatches 36; Indels 50; Gaps 7;
QY 27 HYQTLIDGVRLHSLYTERGPHSDSCGLTLRLRLDPFCPPADQIDRFVQYIDANA 86
DB 411 HYQYL-----SMPDHGVP---SEPGVYL-----SFLDQINQ 438
QY 87 RGEA-----GVGHALFGRGTGMLACYLKE-----RGLAAG-----DAIAEIRLRGCS 132
DB 439 RQESLPHAGPLIVHCASIGRGITIVIMLMENISTGLDCCDIDIOKTQMVRAQRSGM 498
QY 133 IETYEQER-----AVPQFYQRTK 150
DB 499 VQTEHQYKFTIYVAIAQFIETTK 520
RESULT 15
PTN6_MOUSE STANDARD: PRT: 595 AA.
ID PTN6_MOUSE 063872; 063873; Q63874; Q9QVA6; Q9QVA7; Q9QVA8; Q9ROV6;
AC P29351; 063872; 063873; Q63874; Q9QVA6; Q9QVA7; Q9QVA8; Q9ROV6;
AC Q921G3; 035128;
DT 01-DEC-1992 (Ref. 24, Created)
DT 15-JUN-2002 (Ref. 41, Last sequence update)
DT 15-JUN-2002 (Ref. 41, Last annotation update)
DE Protein-tyrosine phosphatase, non-receptor type 6 (EC 3.1.3.48)
DE (Protein-tyrosine phosphatase 1C) (PTP-1C) (Hematopoietic cell
DE protein-tyrosine phosphatase) (70Z-SHP) (SH-PTP1).
GN PTPN6 OR PTP1C OR HCP OR HCPH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DKA/2;
RX MEDLINE=92123209; PubMed=1732748;
RA Yi T., Cleveland J.L., Ihle J.N.;
RT "Protein tyrosine phosphatase containing SH2 domains:
RT characterization, preferential expression in hematopoietic cells, and
RT localization to human chromosome 12p12-p13.";
RL Mol. Cell. Biol. 12:836-846(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92236615; PubMed=1373816;
RA Matthews R.J., Howe D.B., Flores E., Thomas M.L.;
RT "Characterization of hematopoietic intracellular protein tyrosine
RT phosphatases: description of a phosphatase containing an SH2 domain
RT and another enriched in proline-, glutamic acid-, serine-, and
RT threonine-rich sequences.";
RL Mol. Cell. Biol. 12:2396-2405(1992).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANTS MOTHEATEN AND VIABLE MOTHEATEN.
RC STRAIN=C57BL/6J; TISSUE=Bone marrow;
RX MEDLINE=9331972; PubMed=8324828;
RA Schultz L.D., Schweitzer P.A., Rejan T.V., Yi T., Ihle J.N.,
RA Matthews R.J., Thomas M.L., Beier D.R.;
RT "Mutations at the murine motheaten locus are within the hematopoietic
RT cell protein-tyrosine phosphatase (Hcpn) gene.";
RL Cell 73:1445-1454(1993).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
RX MEDLINE=9811780; PubMed=9445485;
RA Ansari-Lari M.A., Oeltyen J.C., Schwartz S., Zhang Z., Muzny D.M.,
RA Lu J., Gorrell J.H., Chinnault A.C., Belmont J.W., Miller W.,

RA Gibbs R.A.;
 RT Comparative sequence analysis of a gene-rich cluster at human
 RT chromosome 12p13 and its syntenic region in mouse chromosome 6.;
 RL Genome Res. 8:29-40(1998).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND SUBUNIT.
 RC STRAIN-C3H; TISSUE-Adrenal gland;
 RX MEDLINE-93348302; PubMed-10419485;
 RA Martin A., Tsui H.W., Shulman M.J., Isenman D., Tsui F.W.;
 RT Murine SHP-1 splice variants with altered Src homology 2 (SH2)
 RT domains. Implications for the SH2-mediated intramolecular regulation
 RT of SHP-1.;
 RL J. Biol. Chem. 274:21725-21734(1999).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RT Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 54-68; 128-135; 137-151; 242-252; 278-285; 293-308 AND
 RP 373-382, AND PHOSPHORYLATION.
 RX MEDLINE-93054686; PubMed-1385421;
 RA Yeung Y.-G., Berg K.L., Pixley F.J., Angeletti R.H., Stanley E.R.;
 RT Protein tyrosine phosphatase-1C is rapidly phosphorylated in tyrosine
 RT in macrophages in response to colony stimulating factor-1.;
 RL J. Biol. Chem. 267:23447-23450(1992).
 CC -1- FUNCTION: PLAYS A KEY ROLE IN HEMATOPOIESIS. THIS PTASE ACTIVITY
 CC MAY DIRECTLY LINK GROWTH FACTOR RECEPTORS AND OTHER SIGNALING
 CC PROTEINS THROUGH PROTEIN-TYROSINE PHOSPHORYLATION. THE SH2 REGIONS
 CC MAY INTERACT WITH OTHER CELLULAR COMPONENTS TO MODULATE ITS OWN
 CC PHOSPHATASE ACTIVITY AGAINST INTERACTING SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SUBUNIT: Monomer.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN HEMATOPOIETIC
 CC CELLS
 CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; are
 CC produced by alternative splicing.
 CC -1- PTM: Phosphorylated on tyrosine residues.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. SUBCLASS THAT CONTAINS SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M68902; AAA37796.1; -;
 DR EMBL; M90389; AAA40007.1; -;
 DR EMBL; S63763; -; NOT_ANNOTATED_CDS.
 DR EMBL; S63764; -; NOT_ANNOTATED_CDS.
 DR EMBL; S63803; -; NOT_ANNOTATED_CDS.
 DR EMBL; AC002397; AAC36009.1; -;
 DR EMBL; AC002397; AAC36008.1; -;
 DR EMBL; U65955; AAD00152.1; -;
 DR EMBL; U65952; AAD00152.1; JOINED.
 DR EMBL; U65953; AAD00152.1; JOINED.
 DR EMBL; U65954; AAD00152.1; JOINED.
 DR EMBL; U65955; AAD00151.1; -;
 DR EMBL; U65951; AAD00151.1; JOINED.
 DR EMBL; U65952; AAD00151.1; JOINED.
 DR EMBL; U65953; AAD00151.1; JOINED.
 DR EMBL; U65954; AAD00151.1; JOINED.
 DR EMBL; BC012660; AAH12660.1; -;
 DR PIR; A44390; A44390.
 DR HSSP; P29350; 1GW2.
 DR MGD; MGI:96055; Hcph.
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR000387; TYR_phosphatase.

DR InterPro; IPR000242; TYR_PP.
 DR Pfam; PF00017; SH2; 2.
 DR Pfam; PF00102; Y_phosphatase; 1.
 DR PRINTS; PRO0700; PRTRYPHATASE.
 DR PRINTS; PRO0401; SH2DOMAIN.
 DR PRODOM; PD000093; SH2; 2.
 DR SMART; SM00194; PTPC; 1.
 DR SMART; SM00252; SH2; 2.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS0056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS0055; TYR_PHOSPHATASE_PTP; 1.
 DR PROSITE; PS0001; SH2; 2.
 KW Hydrolyase; SH2 domain; Repeat; Phosphorylation; Alternative splicing.
 FT DOMAIN 4 100
 FT DOMAIN 110 213
 FT DOMAIN 269 514
 FT ACT_SITE 453 453
 FT VARSPPLIC 1 3
 FT VARSPPLIC 1 39
 FT VARSPPLIC 40 44
 FT VARIANT 77 99
 FT VARIANT 100 595
 FT CONFLICT 240 240
 FT CONFLICT 572 572
 FT CONFLICT 586 586
 SQ SEQUENCE 595 AA; 67559 MW; CF17300D032638D2 CRC64;
 Query Match 13.0%; Score 104; DB 1; Length 595;
 Best Local Similarity 27.5%; Pred. No. 0.0074;
 Matches 39; Conservative 17; Mismatches 36; Indels 50; Gaps 7;
 QY 27 HYGFLDLGVRHLYSLTERKPRHSDSCPGILTHRLAIIPFCPPAPQIDRFQVIVDEANA 86
 DB 411 HYGL-----SWPDHGV-----SEPGVL-----SFLDING 438
 QY 87 RGEA-----VGVCALGFGRTGTMACYLKE-----RGLAAG-----DAIAEIRRLRPGS 132
 DB 439 RQSLRHAGIIVHCAGIGRTITIVIMKESISTKGLDCDIDIKTIQWRAQRSGM 498
 QY 133 IETYEOK-----AVFOFYORTK 150
 DB 499 VQTEAQYKFIYVAIAQFIETTK 520

Search completed: November 4, 2002, 12:47:03
 Job time : 27 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 2002, 12:43:42 ; Search time 43 Seconds
(without alignments)
335.353 Million cell updates/sec

Title: US-09-619-380-2

Perfect score: 798

Sequence: 1 MGVPNFSWVLPGRAGLA.....GSIETYEOKAVQFYQRTK 150

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PTR_73:*

1: ptr1:*

2: ptr2:*

3: ptr3:*

4: ptr4:*

Result No. Score Query Match Length DB ID Description

1	242.5	30.4	190	2	T31661	hypothetical prote
2	186.5	23.4	681	2	E88158	protein C17G10.4c
3	186.5	23.4	708	2	T34097	hypothetical prote
4	186.5	23.4	1063	2	T34097	probable protein-t
5	157.5	19.1	161	2	B90417	conserved hypotet
6	152.5	18.1	151	2	C75157	protein tyrosine p
7	146	18.3	146	2	G71181	hypothetical prote
8	145.5	18.2	537	2	T50099	probable protein-t
9	143	17.9	446	2	T33986	hypothetical prote
10	142.5	17.9	383	2	F84771	hypothetical prote
11	137.5	17.2	551	2	S56283	protein-tyrosine-p
12	133	16.7	177	2	D75309	protein-tyrosine-p
13	125	15.7	376	2	S19740	hypothetical prote
14	119	14.9	168	2	T41891	phosphotyrosine ph
15	117	14.7	165	2	A82246	probable phosphat
16	115	14.4	437	2	D83329	conserved hypotet
17	115	14.4	541	2	JC5937	mRNA guanylyltrans
18	115	14.4	597	2	JC5936	mRNA guanylyltrans
19	114	14.3	168	1	A40781	dual specificity p
20	114	14.3	398	2	T08716	protein-tyrosine-p
21	113	14.2	261	2	T25206	hypothetical prote
22	113	14.2	511	2	JC7198	hypothetical prote
23	113	14.2	585	2	T18885	hypothetical prote
24	113	14.2	598	2	JC7197	protein-tyrosine-p
25	110	13.8	535	2	A46101	protein-tyrosine-p
26	110	13.8	548	2	B46101	protein-tyrosine-p
27	107	13.4	220	2	T10279	protein-tyrosine-p
28	105	13.2	177	2	A12400	hypothetical prote
29	104	13.0	276	2	T48906	protein-tyrosine-p

30	104	13.0	595	1	S20825	protein-tyrosine-p
31	104	13.0	595	1	A44390	protein-tyrosine-p
32	103.5	13.0	159	2	H64326	hypothetical prote
33	99	12.4	190	2	T43172	probable protein-t
34	98.5	12.3	699	2	JC6132	protein-tyrosine-p
35	98	12.3	360	1	JH0692	protein-tyrosine-p
36	98	12.3	700	1	S12053	protein-tyrosine-p
37	98	12.3	2051	2	T30938	receptor tyrosine
38	97.5	12.2	549	2	S53427	protein-tyrosine-p
39	97.5	12.2	656	1	A55574	protein-tyrosine-p
40	97.5	12.2	656	1	UC4263	protein-tyrosine-p
41	96.5	12.1	147	2	C64512	hypothetical prote
42	96.5	12.1	150	2	T21489	hypothetical prote
43	94.5	11.8	173	2	JC5982	prenylated protein
44	94.5	11.8	377	1	A48711	protein-tyrosine-p
45	94	11.8	284	2	T23421	hypothetical prote

ALIGNMENTS

Result 1

T31661 hypothetical protein COS41.7 - sea squirt (Clona intestinalis)

C/Species: Clona intestinalis

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C/Accession: T31661

R/Blind, A.P.; Clark, V.; Jones, S.J.; Lettgeb, S.; Dobson, R.; Tweedle, S.

submitted to the EMBL Data Library, December 1996

A/Reference number: 221049

A/Accession: T31661

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-190 <BIR>

A/Cross-references: EMBL:283760; NID:e1014349; PID:e289997; PIDN:CAB06051.1

C/Genetics:

A/Introns: 56/1; 87/1

Query Match Best Local Similarity 42.7%; Pred. No. 1.5e-16; Matches 61; Conservative 18; Mismatches 39; Indels 25; Gaps 3;

QY	5	PNFNSWVLPGRAGLA	PRLPAYQFLDGVRLVSLFGRGPHSDSCPLTLRLRP	64
DB	66	PSNFTVVDGKWLAA	SPICPTTGG-----VHSTLTK-----NTDAPKL-----	107
QY	65	DECPAPDQIDRFVQIV	DEANANGAVGHCALGFGRTGTMACYLKERGLAAGDAIAE	124
DB	108	-----NOTVMTKSNFK	EFVVAHVAVGANGKIGSTMLACYLKKMRISADALKE	160
QY	125	IRRLRGSIETYEOKA	VQFYQ 147	
DB	161	IRMRPGSVSTEOEKA	VQFYQ 183	

Result 2

E88158 protein C17G10.4c [imported] - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C/Accession: E88158

R/Anonymous, The C. elegans Sequencing Consortium.

A/Title: Genome sequence of the nematode C. elegans: a platform for investigating blo

A/Reference number: A55000; MIMD:99069613; PMID:9851916

A/Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C-

A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;

A/Accession: E88158

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-681 <STO>

A/Cross-references: GB:chr_II; PIDN:AB93458.1; PID:g7731379; GSPDB:GN00020; CESP:C17G

C/Genetics:

```

A:Gene: C17G10.4c
A:Map position: 2

Query Match          23.4%; Score 186.5; DB 2; Length 681;
Best Local Similarity 29.8%; Pred. No. 2,2e-10;
Matches 50; Conservative 30; Mismatches 49; Indels 39; Gaps 7;

OY      3  VQPNFESWVLPGRLAGLALPR-----LPAH-----YQFLDLGVRHLYSL----- 42
DB      192  VENGDFFMIIIPGKILSPGFHNESRENGEYHAAPDYDFYFRENKYSTIVRLNKNYDA 251

OY      43  ---TERPPHSDSQPGILHLRLRIPDFCPAPDOIDRFVQIVDEANAGEAVGYHCALGF 99
DB      252  SKFTKAGFDHVD-----LFFIDGSTPDEIMLKIKRYVD--NTKG-GVAHVCKRAGL 299

OY      100  GRTGTMTACTIVYKERGLAAGDAIAEIRLRPGSI---ETVEQEKAVF 143
DB      300  GRTGTLLACMMKKEYGLTAGECMGMWLRVCRGVSIVGPOQPIYLIKOKF 347

RESULT 3
T34098
Hypothetical protein C17G10.4a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
A:Accession: T34098
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1708 <JOB>
A:Cross-references: EMBL:U28739; PIDN:AAB93457.1; GSPDB:GN00020; CESP:C17G10.4a
A:Experimental source: strain Bristol N2; clone C17G10
C:Genetics:
A:Gene: CESP:C17G10.4a
A:Map position: 2
A:Introns: 19/3; 84/3; 120/3; 541/3; 605/3; 686/3

Query Match          23.4%; Score 186.5; DB 2; Length 708;
Best Local Similarity 29.8%; Pred. No. 2,3e-10;
Matches 50; Conservative 30; Mismatches 49; Indels 39; Gaps 7;

OY      3  VQPNFESWVLPGRLAGLALPR-----LPAH-----YQFLDLGVRHLYSL----- 42
DB      192  VENGDFFMIIIPGKILSPGFHNESRENGEYHAAPDYDFYFRENKYSTIVRLNKNYDA 251

OY      43  ---TERPPHSDSQPGILHLRLRIPDFCPAPDOIDRFVQIVDEANAGEAVGYHCALGF 99
DB      252  SKFTKAGFDHVD-----LFFIDGSTPDEIMLKIKRYVD--NTKG-GVAHVCKRAGL 299

OY      100  GRTGTMTACTIVYKERGLAAGDAIAEIRLRPGSI---ETVEQEKAVF 143
DB      300  GRTGTLLACMMKKEYGLTAGECMGMWLRVCRGVSIVGPOQPIYLIKOKF 347

RESULT 4
T34097
Probable protein-tyrosine-phosphatase (EC 3.1.3.48) CDC14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 01-Dec-2000
A:Accession: T34097
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1063 <JOB>
A:Cross-references: EMBL:U28739; PIDN:AAB93459.1; GSPDB:GN00020; CESP:C17G10.4b

```

[illegible]

RESULT 6

C:Protein: tyrosine phosphatase PAB0279 - Pyrococcus abyssi (strain Orsay)
 C:Species: Pyrococcus abyssi
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
 C:Accession: C75157
 R:Anonymous, Genoscope
 submitted to the EMBL Data Library, July 1999
 A:Description: Pyrococcus abyssi genome sequence: Insights into archaeal chromosome structure
 A:Reference number: A75001
 A:Accession: C75157
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-151 <KAW>
 A:Cross-references: GB:AJ240284; GB:AL096836; NID:g5457730; PIDN:CAB49338.1; PID:el51522
 A:Experimental source: strain Orsay
 C:Genetics:
 A:Gene: PAB0279

RESULT 7

hypothetical protein PH1732 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence,revision 14-Aug-1998 #text,change 21-Jul-2000
C:Accession: G71181
R:Kawabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekine, M.; Ohtuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kishida, N.; Oguchi, D. N. Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon, Pyrococcus horikoshii strain D. A71000, M01D:58344137; PMID:9679194
A:Accession: G71181
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-146 <KMW>
A:Cross-references: GB:P000007; NID:g3236134; PID:BAA30846.1; PID:g3258163
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank C:Gene: PH1732
C:Superfamily: Vhl-type dual specificity phosphoprotein phosphatase homology
F:30-143/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>
Query Match 18.3% Score 146; DB 2; Length 146;
Best Local Similarity 32.5% Pred. No. 3.7e-07;
Matches 27; Conservative 24; Mismatches 32; Indels 0; Gaps 0;

RESULT 8

Probable protein-tyrosine phosphatase CDC14 homolog [imported] - fission yeast (Schizosaccharomyces pombe)

RESULT 9

hypothetical protein T12B3.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T13986
R:Nelson, J., Mohldmann, P., Antoniou, B., Fulton, B.
submitted to the EMBL Data Library, February 1999
A:Description: The sequence of C. elegans cosmid T12B3.
A:Reference number: Z21450
A:Accession: T13986
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-446 <NEE>
A:Cross-references: EMBL:AF125960; PIDN:AD14737.1; GSPDB:GN00022; CESP:T12B3.1
C:Genetics:
C:Experimental source: strain Bristol N2; clone T12B3
A:Gene: CESP:T12B3.1
A:Map position: 4
A:Introns: 21/1; 100/2; 139/1; 228/3; 271/3; 351/3; 391/3

RESULT 10

RESULT 10

F84771
 hypothetical protein At2g35680 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: F84771
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentos, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, I.A.; Shen, M.; Venkatesh, S.E.; Umayam, L.; Talion, L.;
 Nuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MIMD:20083487; PMID:10617197
 A:Accession: F84771
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-383 <STO>
 A:Cross-references: GB:AE002093; NID:94263787; PIDN:AA15447.1; GSPDB:GM00139
 C:Genetics:
 A:Gene: At2g35680
 A:Map position: 2

Query Match 17.9%; Score 142.5; DB 2; Length 383;
 Best Local Similarity 28.8%; Pred. No. 2.5e-06;
 Matches 42; Conservative 31; Mismatches 58; Indels 15; Gaps 6;

QY 10 WVLGRAGLALRLPAHYVFLDLGYRHLVSLFERGPHSDSCGLTLR-----L 61
 DB 122 WFLQFLLG-AVP-FPSDVPQLKELGCVITLNE--PRTLVPS-SLKYSCIDHLVI 175
 QY 62 RPPDFC-PPAPDQIDRFVQIVDEANMGEAVGVCALGFRTGTMACYLKERGLAAGD 120
 DB 176 ATMDYCAPSMEMALICQAVEIHRNASLGKRTTYHCKAGRGSRSTIVCIYLQKHMPEA 235
 QY 121 AIAEIRLRFGSLEYEQEAFQFY 146
 DB 236 AYSYRSIRPRVLLAAQMKRAVEY 261

RESULT 11
 S56283
 protein-tyrosine-phosphatase (EC 3.1.3.48) CDC14 - yeast (Saccharomyces cerevisiae)
 M:Alternate names: cell division control protein CDC14; protein R005; protein YFR028c
 C:Species: Saccharomyces cerevisiae
 C:Date: 02-Sep-1995 #sequence_revision 12-Apr-1996 #text_change 21-Jul-2000
 C:Accession: S56283; A42784; S38416; S59694; S62239; S63833; S69234
 R:Murkani, Y.; Naitou, M.; Hagihara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasanu
 submitted to the EMBL Data Library, May 1995
 A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces ce
 A:Reference number: S56186
 A:Accession: S56283
 A:Molecule type: DNA
 A:Residues: 1-351 <MUR>
 A:Cross-references: EMBL:D50617; NID:9836685; PID:d1009908; PID:g836783; MIPS:YFR028c
 R:Man, J.; Xu, H.; Grunstein, M.
 J. Biol. Chem. 267, 11274-11280, 1992
 A:Title: CDC14 of Saccharomyces cerevisiae. Cloning, sequence analysis, and transcriptio
 A:Reference number: A42784; MIMD:92283835; PMID:1597462
 A:Accession: A42784
 A:Molecule type: DNA
 A:Residues: 'MPY', 114, 'MRISQFQFM', 124-125, 'A', 127, 'G', 131, 'LRKKALL', 139, 'C', 141, 'PSI'
 A:Cross-references: EMBL:M61194; NID:g171182; PID:g171183; EMBL:S37205
 A:Note: sequence extracted from NCBI backbone (NCBIN:104751, NCBIP:104754)
 R:Mal, B.; Lipp, M.
 submitted to the EMBL Data Library, August 1993
 A:Reference number: S38416
 A:Accession: S38416
 A:Molecule type: DNA
 A:Residues: 'M', 101, 'AL', 104, 'LYD', 110, 'RSGMDSASASAV', 118-421 <MAL>
 A:Cross-references: EMBL:X75077; NID:g407517; PID:g407518
 R:Masaki, S.M.S.; Yasushi, M.Y.M.; Akio, T.E.A.T.
 submitted to the EMBL Data Library, June 1995
 A:Description: Dominant mutants alleles of a yeast protein kinase gene CDC15 suppress th
 A:Reference number: S59694
 A:Accession: S59694

A:Molecule type: DNA
 A:Residues: 1-117, 'P', 119-551 <MAS>
 A:Cross-references: EMBL:D55715; NID:g870755; PID:d1010176; PID:g870756
 R:Murkani, Y.
 submitted to the EMBL Data Library, December 1994
 A:Reference number: S62230
 A:Accession: S62239
 A:Molecule type: DNA
 A:Residues: 1-551 <MUR>
 A:Cross-references: EMBL:PD4602; NID:g893419; PID:d1008630; PID:g893423
 R:Kil, T.; Naitou, M.; Hagihara, H.; Abe, M.; Ozawa, M.; Sasanuma, S.I.; Sasanuma, M.
 Yeast 12, 177-190, 1996
 A:Title: Fifteen open reading frames in a 30.8 kb region of the right arm of Chromoso
 A:Reference number: S63830; MIMD:96287654; PMID:8686381
 A:Accession: S63833
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-551 <EKI>
 A:Cross-references: EMBL:D44602; NID:g893419; PID:d1008630; PID:g893423
 R:Shirayama, M.; Matsui, Y.; Toh-e, A.
 Mol. Gen. Genet. 251, 176-185, 1996
 A:Title: Dominant mutant alleles of yeast protein kinase gene CDC15 suppress the Itel
 A:Reference number: S69234; MIMD:96242150; PMID:8668128
 A:Accession: S69234
 A:Molecule type: DNA
 A:Residues: 1-117, 'P', 119-551 <SHI>
 A:Cross-references: EMBL:D5715; NID:g870755; PIDN:BA09533.1; PID:g870756
 C:Genetics:
 A:Gene: SGD:CDK14
 A:Cross-references: SGD:S0001924; MIPS:YFR028c
 A:Map position: 6R
 C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphat
 F:283/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:289/Binding site: substrate phosphate (Arg) #status predicted

Query Match 17.2%; Score 137.5; DB 2; Length 551;
 Best Local Similarity 28.0%; Pred. No. 1.2e-05;
 Matches 30; Conservative 20; Mismatches 34; Indels 23; Gaps 3;

QY 27 HYQFLDLGVRHLVSLFERGPHSDSCGLTLRLRIPDFCPAPDQIDRFVQIVDEANA 86
 DB 238 HFE---DIGIQLDLIEDG-----TCPDLSI-----VKNFVGAETIYK 274
 QY 87 RGEAVGVCALGFRTGTMACYLKERGLAAGDAIEIRLRGSI 133
 DB 275 RGGKIAVHCRAKAGRTGCLIGAHLYTYGFTANECIEFLPIRGWV 321

RESULT 12
 D75309
 protein-tyrosine phosphatase-related protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: D75309
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MIMD:20036896; PMID:10567266
 A:Accession: D75309
 A:Molecule type: DNA
 A:Residues: 1-117 <WHI>
 A:Status: preliminary
 A:Molecule type: DNA
 A:Cross-references: GB:AE002049; GB:AE000513; NID:g6459945; PIDN:AA11705.1; PID:g645
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR2161
 A:Map position: 1

Query Match 16.7%; Score 133; DB 2; Length 177;
 Best Local Similarity 38.2%; Pred. No. 8.8e-06;
 Matches 34; Conservative 15; Mismatches 32; Indels 8; Gaps 3;

QY 67 CFPAPDID---REVQIVDE---ANARGEAVGHGALGFGRTGTMACYLKERGLAAG 119
DB 90 CFIYDQVPSDRARFGELDELTDLDDGNVYVHCRGGLGRAGLTAACLLV-QAGMKPD 148
QY 120 DAIAERIRLRPGSIETFEQEKAVFOYQR 148
DB 149 DAIALVTRKTRGAIENARQEQIFREFAEQ 177

RESULT 13

S19740
hypothetical protein B - Rhodobacter capsulatus
C:Species: Rhodobacter capsulatus
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 12-Feb-1999
C:Accession: S19740
R:Kelly, D.J.
submitted to the EMBL Data Library, February 1992
A:Reference number: S19739
A:Accession: S19740
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <KEI>
A:Cross-references: EMBL:X63974
C:Superfamily: Vhl-type dual specificity phosphoprotein phosphatase homology F.63-182/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <Vhl>

Query Match 15.7%: Score 125; DB 2; Length 376;
Best Local Similarity 30.4%; Pred. No. 0.00013;

Matches 38; Conservative 16; Mismatches 49; Indels 22; Gaps 3;

QY 31 LLDLGRHLVSLTE-----KQPHSDSCPGTLRLRLRIPFCPPAPDQIDR 76
DB 55 IDMGAAHVTLTLEPDELGMKVBDLGTQGR-----AGMDMRLPLIADYSVTPAFEAR 109
QY 77 EV---QIVDEANARGEAVGHGALGFGRTGTMACYLKERGLAAGDAIAEIRLRPGSI 133
DB 110 WQAEGRVIRAAALRAGADVYVHCRGGLGRAGLTAACLLV-DAIAEIRLRPGSI 169
QY 134 ETVYEQ 138
DB 170 ETPAQ 174

RESULT 14

T41891

phosphotyrosine phosphatase orf1 - Bombyx mori nuclear polyhedrosis virus (isolate T3)

C:Species: Bombyx mori nuclear polyhedrosis virus, BmSNPV

A:Variety: isolate T3

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000

C:Accession: T41891

R:Gomi, S.; Majima, K.; Maeda, S.

J. Gen. Virol. 80, 1323-1337, 1999

A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.

A:Reference number: Z22020; MUID:9281911; PMID:10355780

A:Accession: T41891

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-168 <KAM>

A:Cross-references: EMBL:L33180; NID:g3745835; PIDN:AAC63820.1; PTD:g3745973

A:Experimental source: isolate T3

C:Genetics:

A:Note: ptp

C:Superfamily: dual specificity phosphoprotein phosphatase DUSP3; Vhl-type dual specificity

Query Match 14.9%: Score 119; DB 2; Length 168;

Best Local Similarity 32.6%; Pred. No. 0.0002;

Matches 29; Conservative 15; Mismatches 41; Indels 4; Gaps 2;

QY 55 GFLHLRLRIPFCPPAPDQIDREVQIVDEANAR--GEAVGHGALGFGRTGTMACYLTK 112
DB 77 GLYKIKIVPGQTLPSSESTVOEFIDVETFEKPGMLVGVHCRGGLGRAGLTAACLLV 136

QY 113 ERLAAGDAIAEIRLRPGSIETFEQEKAVFOYQR 139
DB 137 TLGIAPOEALINRPEKARGHKIERQNTYQD 165

RESULT 15

AB2246

probable phosphatase VC1070 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: AB2246

R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.

Chardson, D.; Ermolaeva, M.D.; Yamathayan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: AB2035; MUID:20406853; PMID:10952301

A:Accession: AB2246

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-165 <HEI>

A:Cross-references: GB:AE004188; GB:AE003852; NID:g9655530; PIDN:AAF94229.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1070

A:Map position: 1

Query Match 14.7%: Score 117; DB 2; Length 165;

Best Local Similarity 33.3%; Pred. No. 0.00031;

Matches 39; Conservative 15; Mismatches 43; Indels 20; Gaps 5;

QY 35 GVRHLVSLTERGPHSDSCPGTLRLRLRIPFCPPAPDQIDRFVQ-----IVDEANAR 87
DB 58 GVGELPAEVEKA-----GLOMFAPIEDDC--APDA--AAQDMQHSPLAHOALS 105
QY 88 GEAVGHGALGFGRTGTMACYLKERGLAAGDAIAEIRLRPGSIETFEQEKAVFO 144
DB 106 GEKVALLCMGSGSGRTG-LLAHHLLEKGMPLSEITVQVQLRPGAFTEKVVQVQVHQ 161

Search completed: November 4, 2002, 12:49:51
Job time : 45 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 2002, 12:43:02 ; Search time 82 Seconds
(without alignments)
376.915 Million cell updates/sec

Title: US-09-619-380-2

Perfect score: 798
Sequence: 1 MGVPNPFNSWVLPGRLAGLALPRLPANHYOFLDLGVRHLVSLTERGPPHSDSCPGTLTLR 60

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	798	100.0	150	4	Q9BVJ7 homo sapien
2	793	99.4	150	4	Q9NX48
3	725	90.9	144	11	Q9CWA8
4	242.5	30.4	190	5	P91585 clona intes
5	186.5	23.4	681	5	P81300 caenorhabd1
6	186.5	23.4	709	5	P81299 caenorhabd1
7	186.5	23.4	1063	5	Q09976 caenorhabd1
8	170.5	21.4	447	4	Q91600
9	169	21.2	147	1	Q8X270
10	168.5	21.1	459	4	O43183
11	168.5	21.1	471	4	O60730
12	168.5	21.1	498	4	O60729
13	163.5	20.5	195	17	Q8ZX04
14	162.5	20.4	157	17	Q974T2
15	160.5	20.1	383	4	O60728 homo sapien
16	160.5	20.1	594	4	Q9UNH5

17	160.5	20.1	623	4	O60727	O60727 homo sapien
18	157.5	19.7	151	17	Q97VZ7	Q97VZ7 sulfobolus
19	152.5	19.1	161	17	Q9V1L1	Q9V1L1 pyrococcus
20	151	18.9	553	5	Q9V1W7	Q9V1W7 drosophila
21	147.5	18.5	162	17	Q8U0C5	Q8U0C5 pyrococcus
22	147.5	18.5	580	4	O43171	O43171 homo sapien
23	146	18.3	146	17	O59385	O59385 pyrococcus
24	145.5	18.2	537	3	Q9P7H1	Q9P7H1 schizosacch
25	143	17.9	446	5	Q9TAX0	Q9TAX0 caenorhabd1
26	137.5	17.2	326	3	Q05673	Q05673 saccharomyc
27	137.5	17.2	551	3	Q05180	Q05180 saccharomyc
28	137	17.2	245	10	Q940U5	Q940U5 arabidopsi
29	137	17.2	337	10	Q9ZOP1	Q9ZOP1 arabidopsi
30	136	17.0	693	6	Q9N091	Q9N091 macaca fasc
31	133	16.7	177	16	Q9RS63	Q9RS63 delnoccoc
32	133	16.7	177	16	Q9RS63	Q9RS63 delnoccoc
33	128.5	16.1	649	5	Q9VY44	Q9VY44 drosophila
34	128	16.0	362	10	Q9AUS9	Q9AUS9 oryza sativ
35	125	15.7	419	2	O07839	O07839 rhodobacter
36	124.5	15.6	542	3	Q9P8D4	Q9P8D4 candida alb
37	123.5	15.5	752	4	Q9ULE6	Q9ULE6 homo sapien
38	120	15.0	491	12	Q8QUP6	Q8QUP6 infectious
39	119	14.9	168	12	Q92506	Q92506 bombyx mori
40	117	14.7	165	16	Q9KT35	Q9KT35 vibrio chol
41	116	14.5	168	12	Q9DXA4	Q9DXA4 bombyx mori
42	115	14.4	437	16	Q910U5	Q910U5 pseudomonas
43	115	14.4	463	10	Q94JH9	Q94JH9 oryza sativ
44	115	14.4	581	11	Q9DCCL	Q9DCCL mus musculu
45	115	14.4	597	4	Q8WUM8	Q8WUM8 homo sapien

ALIGNMENTS

RESULT 1	Q9BVJ7	PRELIMINARY:	PRT:	150 AA.
AC	Q9BVJ7:			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Hypothetical 16.6 kDa protein.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LUNG;			
RA	Strasbourg R.;			
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC001140; AAH01140.1; -			
DR	InterPro; IPR000340; DS_phosphatase.			
DR	InterPro; IPR000387; TYR_phosphatase.			
DR	Pfam; PF00782; DSPC; 1.			
DR	PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.			
DR	PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.			
KW	Hydrolase; Hypothetical protein.			
SQ	SEQUENCE 150 AA; 1658 MW; 4872EFA0434B1B5F CRC64;			
Query Match	100.0%; Score 798; DB 4; Length 150;			
Best Local Similarity	100.0%; Pred. No. 1.8e-71;			
Matches 150; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 MGVPNPFNSWVLPGRLAGLALPRLPANHYOFLDLGVRHLVSLTERGPPHSDSCPGTLTLR 60			
DB	1 MGVPNPFNSWVLPGRLAGLALPRLPANHYOFLDLGVRHLVSLTERGPPHSDSCPGTLTLR 60			
QY	61 LRLPDCPPAPQIDRFVQIVDEANARGAAGVGHGALGFRGTMTACTYLVKRGIAAD 120			
DB	61 LRLPDCPPAPQIDRFVQIVDEANARGAAGVGHGALGFRGTMTACTYLVKRGIAAD 120			
QY	121 ATAELRLRPGSIETVEDEKAVGFQYQRTK 150			

```

Db 121 AIAEIRRLRGSIETYEQEKAVFOFYORTK 150
|||||
RESULT 2
09CX48 PRELIMINARY; PRT; 150 AA.
AC 09CX48;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE CDNA FLJ20442 f1s, clone KAT04828.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Ohtsuka M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isono T., Sugano S.,
RT "NEBO human cDNA sequencing project."
RL EMBL: AK000449; BAA91172.1;
DR Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSpe; 1.
DR SMART: SM00012; PTPC_DSPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
KW Hydrolase.
SQ SEQUENCE 150 AA; 16598 MW; 4B72EDCD34B1B5F CRC64;

Query Match 99.4%; Score 793; DB 4; Length 150;
Best Local Similarity 99.3%; Pred. No. 5.7e-71;
Matches 149; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVOPPRFSSVNLGRLAGLALPLPAHYQFLDGLGVHLYSLTERGPPHSDSCGTLHRLPF 60
DB 1 MGVOPPRFSSVNLGRLAGLALPLPAHYQFLDGLGVHLYSLTERGPPHSDSCGTLHRLPF 60
QY 61 LRIPDFCPAPDOIDRFVQIVDEANRGEAVGVHCAFGRTGTMLACYLKERGLAAGD 120
DB 61 LRIPDFCPAPDOIDRFVQIVDEANRGEAVGVHCAFGRTGTMLACYLKERGLAAGD 120
QY 121 AIAEIRRLRGSIETYEQEKAVFOFYORTK 150
DB 121 AIAEIRRLRGSIETYEQEKAVFOFYORTK 150

RESULT 3
09CW48 PRELIMINARY; PRT; 144 AA.
AC 09CW48;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE 1300005N15R1k protein (Fragment).
GN 1300005N15R1k.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LIVER;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawada J., Shingawa A., Shidara K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishii K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

```

```

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schmitt L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,
RA Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaldo M.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hune D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaele P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL: AK004912; BAB23663.1;
DR MGI: MGI:1915690; 1300005N15R1k.
DR InterPro: IPR000387; TYR_phosphatase.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
KW Hydrolase.
FT NON_TER
SQ SEQUENCE 144 AA; 16043 MW; ED7BA64AABBD1052 CRC64;

Query Match 90.9%; Score 725; DB 11; Length 144;
Best Local Similarity 94.4%; Pred. No. 3e-64;
Matches 136; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 7 NFSVNLGRLAGLALPLPAHYQFLDGLGVHLYSLTERGPPHSDSCGTLHRLPF 66
DB 1 NFSVNLGRLAGLALPLPAHYQFLDGLGVHLYSLTERGPPHSDSCGTLHRLPF 60
QY 67 CPAPDOIDRFVQIVDEANRGEAVGVHCAFGRTGTMLACYLKERGLAAGDAIEIR 126
DB 61 CPAPDOIDRFVQIVDEANRGEAVGVHCAFGRTGTMLACYLKERGLAAGDAIEIR 120
QY 127 LRPGSIETYEQEKAVFOFYORTK 150
DB 121 LRPGSIETYEQEKAVFOFYORTK 144

RESULT 4
09P385 PRELIMINARY; PRT; 190 AA.
AC 09P385;
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE COS4.7.
OS Clona intestinalis.
OC Eukaryota; Metazoa; Chordata; Trochozoa; Ascidacea; Enterogona;
OC Phlebobranchia; Clonidae; Clona.
OX NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RA Bird A.P., Clark V., Jones S.J.M., Leitgeb S., Dobson R., Tweedie S.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=98208558; PubMed=9539755;
RA Stimm M.W., Leitgeb S., Clark V.H., Jones S.J.M., Bird A.;
RT "Gene number in an invertebrate chordate, Clona intestinalis."
RL Proc. Natl. Acad. Sci. U.S.A. 95:4437-4440(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99148102; PubMed=10024242;
RA Stimm M.W., Leitgeb S., Charlton J., Jones S.J.M., Harris B.R.,
RT "Methylated transposable elements and methylated genes in a
RL chordate genome."
RN [4]
RP SEQUENCE FROM N.A.
RA Stimm M.W., Bird A.;

```

```

RT "Sequence analysis of transposable elements in the Sea Squirt, Clona
TT Internalis";
RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z63760; CAB06051.1; -.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; Tyr.phosphatase.
DR SMART; SMO0012; PTPc_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; UNKNOWN.1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
SQ SEQUENCE 190 AA; 21292 MW; F8934267DDAD27E CRC64;

Query Match          30.4%; Score 242.5; DB 5; Length 190;
Best Local Similarity 42.7%; Pred. No.2.8e-16;
Matches    61; Conservative   18; Mismatches   39; Indels    25; Gaps      3

OY 5 PNFSVLTLEGRAGALPRLPAHYQFLDLGYRHVLSTLERGPSPSDCPGLTLRLRIIP 64
Db 66 PNFTWVPDGMVAACASPCPYGQ-----TWHSYTLKK-----NYDAPEKL----- 107
OY 65 DCPAPADIDFVQIVVDANARGEAVGVCAIGFRGTMTACTYLKRGILAGATAIE 124
Db 108 -----NQTWFTSINFEFGVVQAVAHVCANGISCTIMACTLVKKRISADAKE 160
OY 125 IRRLRPGSIETYEOKKAVQFYQ 147
Db 161 IRRMRPGSVSETEQEKAVEQFYQ 183

RESULT 5
P81300 PRELIMINARY; PRT; 681 AA.
AC P81300;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 76.7 Kda protein C17G10.4C in chromosome II.
GN C17G10.4 OR CDC14.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdittida; Rhabdittidae;
OC Rhabdittidae; Pelodertinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL NZ;
RA Johnson D.;
RN Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
[2]
RP REVISIONS.
RC STRAIN-BRISTOL NZ;
RA Waterston R.;
RN Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Ernsting B.R., Li L., Wishart M.J., Dixon J.E.;
RN Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.
-1 ALTERNATIVE PRODUCTS: THE ORFS C17G10.4A, C17G10.4B AND C17G10.4C
REPRESENT PROBABLE ALTERNATIVE SPLICED FORMS OF A YET HYPOTHETICAL
PROTEIN.
-1 SIMILARITY: SOME TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
TYROSINE PHOSPHATASE FAMILY. CDC14 SUBFAMILY.
EMBL; U28739; AAB93458.1; -.
DR EMBL; AF000363; AAB94407.1; -.
DR WormPeP; C17G10.4C; CE08288.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; Tyr_phosphatase.
DR InterPro; IPR000242; Ty_Pp.
DR Pfam; PF00102; Y_phosphatase; 1.
DR SMART; SMO0012; PTPc_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
KW Hypothetical protein; Alternative splicing.
SEQUENCE 681 AA; 76698 MW; ZC056CD6FA03ADDCC CRC64;
```

```

Query Match 23.4%; Score 186.5; DB 5; Length 681;
Best Local Similarity 29.8%; Pred. No. 4.3e-10;
Matches 50; Conservative 30; Mismatches 49; Indels 39; Gaps 7

Oy 3 VOPNFSWVLPGRLAGLALPR-----LPAH-----YQFLDLGVRHLVSL----- 42
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 192 VENGDFNMIIIPGKILISFCGPHNDSRENGCYPHYAPDYFDYRENNKVSIVIRLNMANNYDA 251
Oy 43 -----TEGPPHSDSCPEQLTLRLRIPFCPPAPDQIDRFQVQIDEANRANGAENGVCALGF 99
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 252 SKFTFAGFDHVD-----LFFIDGSTPSDEIMLKFIVVD--NTKG-GVAVHCAGI 299
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 100 GRTGTMACIYLVKENGLAGDAIAIETRRLRPSGI-----ETYGDEKAVF 143
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 300 GRTGTLIACMMKKEGYLTAGECKMGLVRCRPSGVIQPOOPYLIEOKF 347

RESULT 6
P81299          PRELIMINARY;          PRT:          709 AA.
AC      P81299:
DT      01-JUN-1998 (TREMBLrel. 06, Created)
DT      01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      Hypothetical 79.7 kDa protein C17G10.4A in chromosome II.
GN      C17G10.4.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Pelodietinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
   [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-BRISTOL N2;
RA      Johnson D.;
RL      Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
   (2)
RP      REVISIONS.
RC      STRAIN-BRISTOL N2;
RA      Waterston R.;
RL      Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
CC      -1- ALTERNATIVE PRODUCTS: THE ORFS C17G10.4A, C17G10.4B AND C17G10.4C
CC      REPRESENT PROBABLE ALTERNATIVE SPLICED FORMS OF A YET HYPOTHETICAL
CC      PROTEIN.
CC      -1- SIMILARITY: SOME, TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC      TYROSINE PHOSPHATASE FAMILY. CDC14 SUBFAMILY.
CC      EMBL: U28739; AAB93457.1; --
DR      Wormpep: C17G10.4A; CE16860.
DR      InterPro: IPR000340; DS_phosphatase.
DR      InterPro: IPR000387; TYR_phosphatase.
DR      InterPro: IPR000242; TYR_PP.
DR      Pfam: PF00102; Y_phosphatase; 1.
DR      SMART: SM00012; PTPC_DSPC; 1.
DR      PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR      PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
KW      Hypothetical protein; Alternative splicing.
SQ      SEQUENCE 709 AA; 79739 MW; AB8873DBEC7FEDB0C CRC64;

Query Match 23.4%; Score 186.5; DB 5; Length 709;
Best Local Similarity 29.8%; Pred. No. 4.5e-10;
Matches 50; Conservative 30; Mismatches 49; Indels 39; Gaps 7;

Oy 3 VOPNFSWVLPERLGLALPR-----LPAH-----YQFLDLGVRHLVSL----- 42
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 192 VENGDFNMIIIPKILISFCGPHNDSRENGCYPHYAPDYFDYRENNKVSIVIRLNMANNYDA 251
Oy 43 -----TEGPPHSDSCPEQLTLRLRIPFCPPAPDQIDRFQVQIDEANRANGAENGVCALGF 99
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 252 SKFTFAGFDHVD-----LFFIDGSTPSDEIMLKFIVVD--NTKG-GVAVHCAGI 299
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 100 GRTGTMACIYLVKENGLAGDAIAIETRRLRPSGI-----ETYGDEKAVF 143
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 300 GRTGTLIACMMKKEGYLTAGECKMGLVRCRPSGVIQPOOPYLIEOKF 347

```

```
RESULT 7
ID 009976 PRELIMINARY; PRT; 1063 AA.
AC 009976; 009955;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 120.2 kDa protein.
GN ClpG10.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; Pubmed=9851916;
RA None;
RT "Investigating biology. The C. elegans Sequencing Consortium."
RT Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Johnson D.;
RT "The sequence of C. elegans cosmid ClpG10."
RT Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission."
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: U28739; AAB93459.1; -.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; TYR_PP.
DR Pfam: PF00102; Y_phosphatase; 1.
DR SMART: SM00012; PTPC_DSPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
KW Hydrolase; Hypothetical protein.
SQ SEQUENCE 1063 AA; 120248 MW; 2B0494459F710197 CRC64;

Query Match 23.4%; Score 186.5; DB 5; Length 1063;
Best Local Similarity 29.8%; Pred. No. 7.2e-10;
Matches 50; Conservative 30; Mismatches 49; Indels 39; Gaps 7;

QY 3 VQPPNFSWVLPGRLAGLALP-----LPAH-----YQFLDGVRLHVSLS----- 42
DB 192 VENGDEFWIIPGKILSFCGHNESRENGYYPYHAPDYFFYFRNKVSTYIRLNKKNYDA 251
QY 43 ---TENGPPHSDCPGITLRLRIPDPAPDQIDRFVOIDEANRGEAVGVHCAIGF 99
DB 252 SKFTKAGFDHVD-----LFFIDGSTPSDEIMKFIKIVD--NKG-GVAVHCAGL 299
QY 100 GRGTGMIACLYLVERGIAGDAIAIRLRPGSI---ETFEDEKAVF 143
DB 300 GRGTGLIACWMAKKEYGLTAGECMGLRVCRRPGSVIGPQRYLLEKQRF 347

RESULT 8
ID 09Y600 PRELIMINARY; PRT; 447 AA.
AC 09Y600;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE WUGSC:H_DJ1166619.1 protein.
GN WUGSC:H_DJ1166619.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99063792; Pubmed=9847074;
RA Sulston J.E., Waterston R.;
RT "Toward a complete human genome sequence."
RT Genome Res. 8:1097-1108(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Zidanic M., Wohldman P., Holmes A.;
RT "The sequence of Homo sapiens PAC clone RP5-116619."
RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC006024; AAD15415.1; -.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; TYR_PP.
DR Pfam: PF00102; Y_phosphatase; 1.
DR SMART: SM00012; PTPC_DSPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase.
SQ SEQUENCE 447 AA; 51614 MW; 962ACC30F16D3EF CRC64;

Query Match 21.4%; Score 170.5; DB 4; Length 447;
Best Local Similarity 29.4%; Pred. No. 1e-08;
Matches 50; Conservative 27; Mismatches 62; Indels 31; Gaps 7;

QY 7 NFSWVLPGRLAGLALP-----RLPAHY-----QFLDGVRLHVSLSR-----GPPHS 50
DB 185 DLMNIIRDFRIFGCPHSKRLSEGYHOHPETIYQFKNNVTTIRLNKMTDAKRF 244
QY 51 DSCPGITLRLRIPDPAPDQIDRFVOIDEANRGEAVGVHCAIGFRTGTMACYL 110
DB 245 DA-GFDHDLDFADGSTPDIAIVKRFLLDICE--NABG-AIAYVCKAGLGTGLIACYI 299
QY 111 VKENGLAAGDAIAIRLRPGSITFEDEKAVFO-----FYRTK 150
DB 300 MKHYRMTAETIAWVRICRPGVIGPQOFLVMKQTSIMLEGDYFRORLK 349

RESULT 9
ID 08X270 PRELIMINARY; PRT; 147 AA.
AC 08X270;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Protein tyrosine phosphatase.
GN TK-PTP.
OS Pyrococcus kodakarensis.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus.
OX NCBI_TaxID=69014;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KOD1;
RA Jeon S., Fujiwara S., Takagi M., Fukui K., Imanaka T.;
RT "TK-PTP, protein tyrosine phosphatase from hyperthermophilic archaeon Thermococcus kodakarensis KOD1: Enzymatic characteristics and isolation of its substrate proteins."
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB051898; BAB83049.1; -.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; TYR_PP.
DR Pfam: PF00782; DSPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
```



```

QY 7 NFSWVLEGRGLALP-----RLPAHY-----QELLDGVNHLVSLTER---GPPHS 50
DB 215 DLNMIIDRFRTAFCGPSRRLAESGTYQHSPETRYIOYFKNNVYTIIRLNRMADAKRT 274
QY 51 DSCGFLIHLRLIPDFCPAPDQIDRFVQIVDEANRGEAVGVCALGFRTGTMACYL 110
DB 275 DA--GPHHDLFFADGSPIDAIYKEFLIDCE--NAEG-AIAYVCKAGLGTGTLIACYI 329
QY 111 VKERGLAAGDAIEIRLRPGSIEFYOEKAVFO 144
DB 330 MKHYMTAAETIAWYRICRPGSYIGPOOQFLVMK 363

RESULT 13
Q8ZX04 PRELIMINARY; PRT; 195 AA.
AC 08ZX04;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Conserved protein (possible dual specificity phosphatase).
GN PAE1536.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_Taxid=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL: AE009826; AAL6345.1; -.
DR InterPro: IPR000387; TYR_phosphatase.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
KW Complete proteome.
SQ SEQUENCE 195 AA; 22170 MW; DF2648D5987A8D41 CRC64;

Query Match 20.5%; Score 163.5; DB 17; Length 195;
Best Local Similarity 34.0%; Pred. No. 1.9e-08;
Matches 50; Conservative 21; Mismatches 59; Indels 17; Gaps 5;

QY 10 WVLPGRLAGLALPRLPAHYQFLDLDGVNHLVSLTERGPPHSDSCPGTLHLRLIP 64
DB 40 WIRP-RIAGSCMG-RODIEKMAELGVKYVSLAEANEIEYRGWGLLEKTKLMKGMK 97
QY 64 -----PDCFPAPDQIDRFVQIVDEANRGEAVGVCALGFRTGTMACYLKERGLA 117
DB 98 WIRPDPDGPFR--KIDELVEYLKEEAKGSYV-VHCVGIGIRTPVLAAYLATKMK 154
QY 118 AGDAIAEIRLRPGSIEFYOEKAVFO 144
DB 155 ADDAIRVERVNPVSLTDQYVALL 181

RESULT 14
Q974T2 PRELIMINARY; PRT; 157 AA.
AC 0974T2;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical protein ST0580.
GN ST0580.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_Taxid=111955;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Aokai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL: AP000383; BAB5573.1; -.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00102; Y_phosphatase; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50055; TYR_PHOSPHATASE_PP; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 157 AA; 18569 MW; B77C0A14DAB8B061 CRC64;

```

```

Query Match 20.4%; Score 162.5; DB 17; Length 157;
Best Local Similarity 33.8%; Pred. No. 1.9e-08;
Matches 49; Conservative 24; Mismatches 59; Indels 13; Gaps 5;

```

```

QY 10 WVLPGRLAGLALPRLPAHYQFLDLDGVNHLVSLTERGPPHSDSCPGTLHLRLIP 64
DB 4 WIRGIIGSPITPTEDELDMKREGVKRILLPEEW--ELEWAGSMDYFSLIREKGF 61
QY 65 DFC-PAPD-----QIDRFVQIVDEANRGEAVGVCALGFRTGTMACYLKERGLA 119
DB 62 EFLHEPIPDGAPFPDQFLRIY-EMTKGYANLVHCVGIGRTGTIAGYLMFEDDSG 120
QY 120 DAIAEIRLRPGSIEFYOEKAVFO 144
DB 121 EALIEVRYNPGAVQYEQLELLO 145

RESULT 15
Q60728 PRELIMINARY; PRT; 383 AA.
AC 060728;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Cdc14A3 phosphatase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Hao L., Baskerville C., Charbonneau H.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF064103; AAC16660.1; -.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR00242; TYR_PP.
DR Pfam: PF00102; Y_phosphatase; 1.
DR SMART: SM00012; PRPC_DSPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase.
SQ SEQUENCE 383 AA; 43908 MW; 28295D04793D00B7 CRC64;

```

```

Query Match 20.1%; Score 160.5; DB 4; Length 383;
Best Local Similarity 29.9%; Pred. No. 8.4e-08;
Matches 46; Conservative 22; Mismatches 51; Indels 35; Gaps 5;

QY 3 VQPNFWSVLEGRGLA-----GLALPRLPAHYQFLDLDGVNHLVSL----- 42

```

```

Db 175 VENGDENMIVPGKFLAFSGPHPSKXIENGYPPLHAPFAYPFPYFKKHNTTAVVRLNKKIYEA 234
QY 43 ---TERGPPHSDSCPGTLTLHLRLIPDECPAPQIDRFQIYDEANARGENAVGHCALGF 99
Db 235 KRFTDAGFEHYD-----LFFIDGSTPSDNIVRRFNICE--NTEG-AIAVHCKAGL 282
QY 100 GRTGTMACYLVKERGIAGDAIAEIRLRPGSI 133
Db 283 GRTGTLIACYVMKHYRPTHAEIILAMIRICRPGSI 316

```

Search completed: November 4, 2002, 12:48:47
 Job time : 84 secs

FT Domain 74.125
FT /note= "Tyrosine specific protein phosphatase
FT domain signature"
FT Modified-site 41
FT /note= "O-phosphorylated"
FT Modified-site 43
FT /note= "O-phosphorylated"
FT Modified-site 7
FT /note= "N-glycosylated"
XX WO200120004-A2.
XX
XX
XX PD 22-MAR-2001.
XX
XX PF 14-SEP-2000; 2000WO-US25515.
XX
XX PR 15-SEP-1999; 99US-0154141.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y,
PI Lu DM;
XX
XX WPI: 2001-244811/25.
DR N-PSDB: AAF30482.
XX
XX Novel human protein phosphatase and kinase proteins for diagnosis,
PT treatment and prevention of gastrointestinal, immune system,
PT neurological and cell proliferative disorders -
XX
XX Claim 1; Page 91; 103pp; English.
XX
XX The present sequence is that of novel human protein phosphatase and
CC kinase protein PPHK-7, as predicted from Incyte Clone ID NO.
CC 1606974CB1 (see AAF30482). Tissues that express PPHK-7 (as a
CC fraction of total tissues expressing PPHK-7) include reproductive
CC (0.372), cardiovascular (0.140) and endocrine (0.093). Diseases or
CC conditions associated with tissues expressing PPHK-7 (as a
CC fraction of total tissues expressing PPHK-7) include cancer
CC (0.558), inflammation or trauma (0.233) or cell proliferation
CC (0.209). The encoded protein shows homology to Clona Interstinalis
CC tyrosine phosphatase. The invention provides human PPHK-1 to -11
CC polypeptides (see AAB20322-32) and polynucleotides (see AAF30476-86).
CC It also provides expression vectors, host cells, antibodies, agonists
CC and antagonists, as well as methods for diagnosing, treating or
CC preventing disorders associated with expression of PPHK, including
CC gastrointestinal disorders, immune system disorders, neurological
CC disorders and cell proliferative disorders, including cancer.
XX
XX Sequence 150 AA;
SQ
Query Match 100.0%; Score 798; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 4.2e-85;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGVPNPNSSWLPGRLAGLALPRLPAHYQFLDGLGRHLVSTERGPHSDSCPGLTLHR 60
DB 1 MGVPNPNSSWLPGRLAGLALPRLPAHYQFLDGLGRHLVSTERGPHSDSCPGLTLHR 60
QY 61 LRIDPFCPPADQIDRFVQIYDEANARGEAVGHCALGFGRTGTMACYLVKERGLAAGD 120
DB 61 LRIDPFCPPADQIDRFVQIYDEANARGEAVGHCALGFGRTGTMACYLVKERGLAAGD 120
QY 121 AIAEIRRLRPGSIETYEOKAVFOFYQRTK 150
DB 121 AIAEIRRLRPGSIETYEOKAVFOFYQRTK 150
RESULT 2
AAB35275
ID AAB35275 standard; Protein: 150 AA.
XX
AC AAB35275;

XX 08-MAY-2001 (first entry)
DT
XX
XX Human dual specificity phosphatase DSP-11.
DE
XX
XX Human; dual specificity phosphatase, DSP-11; Duchenne muscular dystrophy;
KW MAP-kinase signalling cascade; cell proliferation; cancer; allergy;
KW graft-versus-host disease; autoimmune disease; metabolic disease.
XX
XX Homo sapiens.
OS
XX
XX WO200105983-A1.
XX
XX PD 25-JAN-2001.
XX
XX PF 19-JUL-2000; 2000WO-US19710.
XX
XX PR 20-JUL-1999; 99US-0144557.
XX
XX (CEPT-) CEPTYR INC.
XX
XX Luche RM, Wel B;
PI
XX
XX WPI: 2001-147348/15.
DR N-PSDB: AAF27958.
XX
XX Novel dual specificity phosphatase-11 dephosphorylating activated
PT mitogen-activated protein kinase. Is used to identify agents inhibiting
PT the enzyme activity and modulate cell proliferation -
XX
XX Claim 1; Fig 2; 65pp; English.
XX
XX The present invention provides the protein and coding sequences of the
CC human dual-specificity phosphatase DSP-11. The protein is involved in the
CC MAP-kinase signalling cascade. The sequences can be used in the diagnosis
CC and treatment of Duchenne muscular dystrophy, cancer, graft-versus-host
CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal
CC cell growth and abnormal cell proliferation. The present sequence is the
CC human DSP-11 protein.
XX
XX Sequence 150 AA;
SQ
Query Match 100.0%; Score 798; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 4.2e-85;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGVPNPNSSWLPGRLAGLALPRLPAHYQFLDGLGRHLVSTERGPHSDSCPGLTLHR 60
DB 1 MGVPNPNSSWLPGRLAGLALPRLPAHYQFLDGLGRHLVSTERGPHSDSCPGLTLHR 60
QY 61 LRIDPFCPPADQIDRFVQIYDEANARGEAVGHCALGFGRTGTMACYLVKERGLAAGD 120
DB 61 LRIDPFCPPADQIDRFVQIYDEANARGEAVGHCALGFGRTGTMACYLVKERGLAAGD 120
QY 121 AIAEIRRLRPGSIETYEOKAVFOFYQRTK 150
DB 121 AIAEIRRLRPGSIETYEOKAVFOFYQRTK 150
RESULT 3
ABP43078
ID ABP43078 standard; Protein: 179 AA.
XX
XX ABP43078;
XX
XX 22-AUG-2002 (first entry)
XX
XX Human ovarian antigen HVCAJ27, SEQ ID NO:4210.
DE
XX
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;

KM inflammatory condition; immune disorder; blood disorder;
KM cardiovascular disorder; respiratory disorder; neurological disorder;
KM gastrointestinal disorder; urinary system disorder; drug screening;
KM gene therapy; chromosome mapping; forensic analysis;
KM antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KM antiinflammatory; gynaecological; reproductive.
OS Homo sapiens.
PN WO200200677-A1.
XX
PD 03-JAN-2002.
XX
PE 07-JUN-2001; 2001WO-US18569.
XX
PR 07-JUN-2000; 2000US-209467P.
XX
PA (HOMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI: 2002-147878/19.
XX
N-PSDB; ABQ56155.
XX
PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -
XX
PS Claim 11; SEQ ID NO 4210; 2922pp; English.
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP4328) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen.
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 179 AA:
Query Match 100.0%; Score 798; DB 23; Length 179;
Best Local Similarity 100.0%; Pred. No. 5.2e-85;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MGVQPNFESWVLPGRLAGLALPRLPAHYOFLDLGVRLVSLTERGPPHSDSCPGTLTLR 60
DB 30 MGVQPNFESWVLPGRLAGLALPRLPAHYOFLDLGVRLVSLTERGPPHSDSCPGTLTLR 89
OY 61 LRIPEFCPPAPQIDRFVQIVDEANARGEAVGVHICALGFGRTGTMLACTLVKRGILAAGD 120
DB 61 LRIPEFCPPAPQIDRFVQIVDEANARGEAVGVHICALGFGRTGTMLACTLVKRGILAAGD 120

DB 90 LRIPEFCPPAPQIDRFVQIVDEANARGEAVGVHICALGFGRTGTMLACTLVKRGILAAGD 149
OY 121 AIAEIRLRPGSIETYEOKAVFOYORTK 150
DB 150 AIAEIRLRPGSIETYEOKAVFOYORTK 179
RESULT 4
AAB73231
ID AAB73231 standard; Protein; 150 AA.
XX
XX AAB73231;
XX
XX 11-MAY-2001 (first entry)
XX
XX Human phosphatase BAA91172_h.
DE
XX Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
XX cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
XX congenital muscle disorder; Papillon-Leleuve syndrome; Cowden disease;
XX Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
XX schizophrenia; hamartoma.
OS Homo sapiens.
XX
XX WO200112819-A2.
XX
XX 22-FEB-2001.
XX
XX 11-AUG-2000; 2000WO-US22158.
XX
XX 13-AUG-1999; 99US-0149005.
XX
XX (SUGEN-) SUGEN INC.
XX
XX PLOWMAN GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;
XX WPI: 2001-211226/21.
XX N-PSDB; AAF63563.
XX
XX New protein phosphatase polypeptide for diagnosing and treating
XX phosphatase related disorders such as cancer, schizophrenia, cardiac
XX dysfunction and/or vascular disorders -
XX
PS Claim 6; Fig 5; 138pp; English.
XX
XX The present invention relates to phosphatase proteins and coding
XX sequences. The present sequence is one such phosphatase. Phosphatases are
XX enzymes that catalyse the dephosphorylation of proteins modified by
XX phosphorylation of serine, threonine or tyrosine residues. The
XX phosphatases are useful for treating a variety of diseases: for example
XX cancer e.g. breast, urogenital, prostate, head, neck, lung cancers,
XX synovial sarcomas, renal cell carcinoma, non-small cell lung cancer,
XX hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,
XX glioblastoma, colorectal cancer and thyroid cancer; pathophysiological
XX hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,
XX congenital muscle disorders, Papillon-Leleuve syndrome, Cowden disease,
XX ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan
XX Zonana syndrome, schizophrenia and hamartomas.
XX
SQ Sequence 150 AA:
Query Match 99.4%; Score 793; DB 22; Length 150;
Best Local Similarity 99.3%; Pred. No. 1.6e-84;
Matches 149; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MGVQPNFESWVLPGRLAGLALPRLPAHYOFLDLGVRLVSLTERGPPHSDSCPGTLTLR 60
DB 1 MGVQPNFESWVLPGRLAGLALPRLPAHYOFLDLGVRLVSLTERGPPHSDSCPGTLTLR 60
OY 61 LRIPEFCPPAPQIDRFVQIVDEANARGEAVGVHICALGFGRTGTMLACTLVKRGILAAGD 120
DB 61 LRIPEFCPPAPQIDRFVQIVDEANARGEAVGVHICALGFGRTGTMLACTLVKRGILAAGD 120

OY 121 AIAEIRRLRPGSIETFEQEKAVFQYQRTK 150
 |||||
 DB 121 AIAEIRRLRPGSIETFEQEKAVFQYQRTK 150

RESULT 5

AG67455
 ID AAG67455 standard; Protein; 150 AA.

AC AAG67455;

DT 26-NOV-2001 (first entry)

DE Amino acid sequence of a human polypeptide.

KM Human; protein kinase; protein phosphatase; signal transduction;
 intracellular signalling pathway.

OS Homo sapiens.

PN WO200109345-A1.

PD 08-FEB-2001.

PF 28-JUL-2000; 2000WO-JP05060.

PR 29-JUL-1999; 99JP-0248036.

PR 18-OCT-1999; 99US-0159590.

PR 11-JAN-2000; 2000JP-0118776.

PR 17-FEB-2000; 2000US-0183322.

PR 02-MAY-2000; 2000JP-0183767.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;
 Seno C, Nezu J;

PI WPI; 2001-564736/63.

DR WPI; 2001-564736/63.

XX New genes encoding protein kinase and protein phosphatase, useful for
 identifying modulators which can be used to treat human or animal
 disorders associated with the expression or function of these enzymes -

PS Example 4; Page 297-298; 336pp; Japanese.

XX The specification describes human protein kinase/protein phosphatases.

CC The polypeptides are expected to participate in signal transduction

CC in cells. The kinase phosphatases are connected with intracellular

CC signalling pathways. Antisense oligonucleotides and compounds

CC identified by screening (agonists or antagonists) can be used to

CC treat human or animal disorders associated with the expression

CC or function of the protein. In addition, the polypeptides may be used

CC as target molecules for drug development. The present sequence

CC represents a polypeptide, used in the course of the invention.

XX Sequence 150 AA;

XX Query Match 99.4%; Score 793; DB 22; Length 150;

XX Best Local Similarity 99.3%; Pred. No. 1.6e-84;

XX Matches 149; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 1 MGVPNPNFNSWLPGRLAGLALPRLPAHQFLLDGLVRLVSLTERGPHSDSCPGTLAR 60

DB 1 MGVPNPNFNSWLPGRLAGLALPRLPAHQFLLDGLVRLVSLTERGPHSDSCPGTLAR 60

OY 61 LRIPDFCPAPDQIDRFQYQYDEANRGEAVGHCALFGRTGTMLACYLVKERGLAAGD 120

DB 61 LRIPDFCPAPDQIDRFQYQYDEANRGEAVGHCALFGRTGTMLACYLVKERGLAAGD 120

OY 121 AIAEIRRLRPGSIETFEQEKAVFQYQRTK 150

DB 121 AIAEIRRLRPGSIETFEQEKAVFQYQRTK 150

DB 121 AIAEIRRLRPGSIETFEQEKAVFQYQRTK 150

RESULT 6

AG67634
 ID AAG67634 standard; Protein; 150 AA.

AC AAG67634;

DT 26-NOV-2001 (first entry)

DE Amino acid sequence of a human protein.

KM Human; protein kinase; protein phosphatase; signal transduction.

OS Homo sapiens.

PN WO200109316-A1.

PD 08-FEB-2001.

PF 28-JUL-2000; 2000WO-JP05061.

PR 29-JUL-1999; 99JP-0248036.

PR 18-OCT-1999; 99US-0159590.

PR 11-JAN-2000; 2000JP-0118776.

PR 17-FEB-2000; 2000US-0183322.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;
 Seno C, Nezu J;

PI WPI; 2001-570286/64.

DR WPI; 2001-570286/64.

XX New genes encoding proteins with protein kinase/protein phosphatase

XX activity, useful in the diagnosis and treatment of diseases -

PS Example 4; Page 197-198; 233pp; Japanese.

XX The specification describes human protein kinase/protein phosphatases.

CC It is expected that the protein kinase/protein phosphatase gene

CC participates in signal transduction in cells. The protein

CC kinase/protein phosphatase polypeptides and polynucleotides are

CC useful for developing diagnostics and treatment agents for human

CC and animal diseases. The protein kinase/protein phosphatase polypeptides

CC are useful as target molecules in designing novel drugs. The protein

CC kinase/protein phosphatase polynucleotides are useful as a source of

CC probes and primers, which may be used to isolate homologous sequences.

CC The present sequence represents a human protein, which is used in the

CC course of the invention.

XX Sequence 150 AA;

XX Query Match 99.4%; Score 793; DB 22; Length 150;

XX Best Local Similarity 99.3%; Pred. No. 1.6e-84;

XX Matches 149; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MGVPNPNFNSWLPGRLAGLALPRLPAHQFLLDGLVRLVSLTERGPHSDSCPGTLAR 60

DB 1 MGVPNPNFNSWLPGRLAGLALPRLPAHQFLLDGLVRLVSLTERGPHSDSCPGTLAR 60

OY 61 LRIPDFCPAPDQIDRFQYQYDEANRGEAVGHCALFGRTGTMLACYLVKERGLAAGD 120

DB 61 LRIPDFCPAPDQIDRFQYQYDEANRGEAVGHCALFGRTGTMLACYLVKERGLAAGD 120

OY 121 AIAEIRRLRPGSIETFEQEKAVFQYQRTK 150

DB 121 AIAEIRRLRPGSIETFEQEKAVFQYQRTK 150

```

RESULT 7
ABB07846
ID ABB07846 standard; protein; 150 AA.
XX
XX ABB07846;
AC
XX
XX 03-JUL-2002 (first entry)
DE Human tyrosine phosphatase-like enzyme.
XX
XX Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cyostatic;
XX cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
XX cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
XX antirhythmic; hypotensive; vulnerrary; gene therapy; human.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX 92.108
XX Region /note="the encoding nucleotide fragment for the above
XX residues is not indicated in the corresponding
XX DNA sequence"
XX
XX WO200220747-A2.
XX
XX 14-MAR-2002.
XX
XX 05-SEP-2001; 2001WO-EP10205.
XX
XX 11-SEP-2000; 2000US-231568P.
XX
XX 06-DEC-2000; 2000US-251403P.
XX
XX (FARB ) BAYER AG.
XX
XX Kossida S;
XX
XX WPI; 2002-339803/37.
XX
XX N-PSDB; ABL40806.
XX
XX New human tyrosine phosphatase-like enzyme polypeptide, regulators of
XX which are useful for preventing, treating diabetes, obesity, cancer,
XX cardiovascular and pulmonary diseases
XX
XX Claim 18; Fig 2; 117pp; English.
XX
XX The invention relates to a purified human tyrosine phosphatase-like
XX enzyme polypeptide. The enzyme can be expressed by standard recombinant
XX methodology. The tyrosine phosphatase-like enzyme and encoding
XX polynucleotides are useful for screening for modulators which are used
XX for treating a tyrosine phosphatase-like enzyme dysfunction related
XX disease such as a central nervous system (CNS) disorder, diabetes,
XX obesity, chronic obstructive pulmonary disease, cardiovascular disease,
XX cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
XX other eating disorder such as bulimia. CNS disorders include brain
XX injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
XX Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
XX Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
XX immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
XX and cardiovascular diseases include myocardial infarction, ischaemic
XX diseases of the heart, atrial and ventricular arrhythmia, hypertensive
XX vascular diseases and peripheral vascular diseases. The enzyme is useful
XX in diagnostic assays for detecting diseases and abnormalities or
XX susceptibility to diseases or abnormalities related to the presence of
XX mutations in the encoding nucleic acid sequences. The present sequence
XX represents the human tyrosine phosphatase-like enzyme polypeptide.
XX
XX Sequence 150 AA:

```

```

Query Match 99.4%; Score 793; DB 23; Length 150;
Best Local Similarity 99.3%; Pred. No. 1,6e-84;
Matches 149; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MGVPPEPNSWVLPGRLAGLALPRLPAAHYOFLLDGVRIHIVSLTERGPHSDSCGTLHR 60
DB 1 MGVPPEPNSWVLPGRLAGLALPRLPAAHYOFLLDGVRIHIVSLTERGPHSDSCGTLHR 60
QY 61 LRIPDFCPAPDDQIDRFVQIVDEANANGEAVGACAFGRTGFMACYLKKEGLAAGD 120
DB 61 LRIPDFCPAPDDQIDRFVQIVDEANANGEAVGACAFGRTGFMACYLKKEGLAAGD 120
QY 121 AIAEIRRLRPGSIETTYEKEAVQFYORTK 150
DB 121 AIAEIRRLRPGSIETTYEKEAVQFYORTK 150

RESULT 8
AAB73211
ID AAB73211 standard; Protein; 150 AA.
XX
XX AAB73211;
AC
XX
XX 11-MAY-2001 (first entry)
DE Murine phosphatase AAO23073_m.
XX
XX Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
XX cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
XX congenital muscle disorder; Papillon-Lejeune syndrome; Cowden disease;
XX Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
XX schizophrenia; hamartoma.
XX
XX Mus sp.
XX
XX WO200112819-A2.
XX
XX 22-FEB-2001.
XX
XX 11-AUG-2000; 2000WO-US22158.
XX
XX 13-AUG-1999; 99US-0149005.
XX
XX (SUGB-) SUGEN INC.
XX
XX Plowman GD, Martinez R, Whyte D, Hall R, Flanagan P, Lloblin M;
XX
XX WPI: 2001-211226/21.
XX
XX N-PSDB; AAF63563.
XX
XX New protein phosphatase polypeptide for diagnosing and treating
XX phosphatase related disorders such as cancer, schizophrenia, cardiac
XX dysfunction and/or vascular disorders
XX
XX Claim 6; Fig 5; 138pp; English.
XX
XX The present invention relates to phosphatase proteins and coding
XX sequences. The present sequence is one such phosphatase. Phosphatases are
XX enzymes that catalyse the dephosphorylation of proteins modified by
XX phosphorylation of serine, threonine or tyrosine residues. The
XX phosphatases are useful for treating a variety of diseases: for example
XX cancer e.g. breast, urogenital, prostate, head, neck, lung cancers,
XX synovial sarcomas, renal cell carcinoma, non-small cell lung cancer,
XX hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,
XX glioblastoma, colorectal cancer and thyroid cancer, pathophysiological
XX hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,
XX congenital muscle disorders, Papillon-Lejeune syndrome, Cowden disease,
XX ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan
XX Zonana syndrome, schizophrenia and hamartomas.
XX
XX Sequence 150 AA:

```

```

Query Match 96.4%; Score 769; DB 22; Length 150;
Best Local Similarity 95.3%; Pred. No. 1e-81;
Matches 143; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 1 MGVPPEPNSWVLPGRLAGLALPRLPAAHYOFLLDGVRIHIVSLTERGPHSDSCGTLHR 60

```

```

Db      121 AIAEIRRLRPGSIETTYEQEKAIVFOFYQRTK 150
      1 MGVOQPNFNSVTLPGRLAGLALPRLPAHYQFLDDGVRAHLSLTERGPHSDSCGTLTLHR 60
      61 LRIPDFCPAPDQIDRFVQIVDEANARGAVGHCALGFGRGTGMLACYLTKERGLAAGD 120
      61 MRIPDFCPAPDQIDRFVQIVDEANARGAVGHCALGFGRGTGMLACYLTKERGLAAGD 120
      121 AIAEIRRLRPGSIETTYEQEKAIVFOFYQRTK 150
      121 AIAEIRRLRPGSIETTYEQEKAIVFOFYQRTK 150

RESULT 9
AAB35276
ID AAB35276 standard; Protein; 150 AA.
AC AAB35276;
XX
DT 08-MAY-2001 (first entry)
DE Murine dual specificity phosphatase DSP-11.
XX
XX Mouse; dual specificity phosphatase; DSP-11; Duchenne muscular dystrophy;
KM MAP-kinase signalling cascade; cell proliferation; cancer; allergy;
KM graft-versus-host disease; autoimmune disease; metabolic disease.
XX
OS Mus sp.
XX
XX WO200105983-A1.
XX
XX 25-JAN-2001.
XX
XX 19-JUL-2000; 2000WO-0519710.
XX
XX 20-JUL-1999; 99US-0144557.
XX
XX (CEPT-) CEPYR INC.
XX
XX Lucbe RM, Wei B;
XX
XX WPI; 2001-147348/15.
XX
XX N-PSDB; AAF27959.
XX
XX Novel dual specificity phosphatase-11 dephosphorylating activated
PT mitogen-activated protein kinase, is used to identify agents inhibiting
PT the enzyme activity and modulate cell proliferation -
XX
XX Example 3; Fig 6; 65pp; English.
XX
XX The present invention provides the protein and coding sequences of the
CC human dual specificity phosphatase DSP-11. The protein is involved in the
CC MAP-kinase signalling cascade. The sequences can be used in the diagnosis
CC and treatment of Duchenne muscular dystrophy, cancer, graft-versus-host
CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal
CC cell growth and abnormal cell proliferation. The present sequence is the
CC mouse DSP-11 protein.
XX
XX Sequence 150 AA;
XX
Query Match 95.7%; Score 764; DB 22; Length 150;
Best Local Similarity 94.7%; Pred. NO. 3.9e-81;
Matches 142; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
XX
1 MGVOQPNFNSVTLPGRLAGLALPRLPAHYQFLDDGVRAHLSLTERGPHSDSCGTLTLHR 60
1 MGVOQPNFNSVTLPGRLAGLALPRLPAHYQFLDDGVRAHLSLTERGPHSDSCGTLTLHR 60
61 LRIPDFCPAPDQIDRFVQIVDEANARGAVGHCALGFGRGTGMLACYLTKERGLAAGD 120
61 LRIPDFCPAPDQIDRFVQIVDEANARGAVGHCALGFGRGTGMLACYLTKERGLAAGD 120
61 MRIPDFCPAPDQIDRFVQIVDEANARGAVGHCALGFGRGTGMLACYLTKERGLAAGD 120
121 AIAEIRRLRPGSIETTYEQEKAIVFOFYQRTK 150
121 AIAEIRRLRPGSIETTYEQEKAIVFOFYQRTK 150

```

```

Db      121 AIAEIRRLRPGSIETTYEQEKAIVFOFYQRTK 150
      121 AIAEIRRLRPGSIETTYEQEKAIVFOFYQRTK 150

RESULT 10
AAB07847
ID AAB07847 standard; Protein; 681 AA.
AC AAB07847;
XX
XX 03-JUL-2002 (first entry)
DE C. elegans protein identified by EMBL Accession No. AF000363.
XX
XX Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytostatic;
KM cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
KM cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
KM antirhythmic; hypotensive; vulnerery.
XX
XX Caenorhabditis elegans.
XX
XX WO200220747-A2.
XX
XX 14-MAR-2002.
XX
XX 05-SEP-2001; 2001WO-EP10205.
XX
XX 11-SEP-2000; 2000US-231568P.
XX
XX 06-DEC-2000; 2000US-251403P.
XX
XX (FARB ) BAYER AG.
XX
XX Kossida S;
XX
XX WPI; 2002-339803/37.
XX
XX New human tyrosine phosphatase-like enzyme polypeptide, regulators of
PT which are useful for preventing, treating diabetes, obesity, cancer,
PT cardiovascular and pulmonary diseases -
XX
XX Disclosure; Fig 3; 117pp; English.
XX
XX The invention relates to a purified human tyrosine phosphatase-like
CC enzyme polypeptide. The enzyme can be expressed by standard recombinant
CC methodology. The tyrosine phosphatase-like enzyme and encoding
CC polynucleotides are useful for screening for modulators which are used
CC for treating a tyrosine phosphatase-like enzyme dysfunction related
CC disease such as a central nervous system (CNS) disorder, diabetes,
CC obesity, chronic obstructive pulmonary disease, cardiovascular disease,
CC cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
CC other eating disorder such as bulimia. CNS disorders include brain
CC injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
CC Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
CC Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
CC and cardiovascular diseases include myocardial infarction, ischemic
CC diseases of the heart, atrial and ventricular arrhythmia, hypertensive
CC vascular diseases and peripheral vascular diseases. The enzyme is useful
CC in diagnostic assays for detecting diseases and abnormalities or
CC susceptibility to diseases or abnormalities related to the presence of
CC mutations in the encoding nucleic acid sequences. The present sequence
CC represents a C. elegans protein identified by EMBL Accn No. AF000363,
CC used for alignment studies with the human tyrosine phosphatase-like
CC enzyme polypeptide.
XX
XX Sequence 681 AA;
XX
Query Match 23.4%; Score 186.5; DB 23; Length 681;
Best Local Similarity 29.8%; Pred. NO. 9.1e-13;
Matches 50; Conservative 30; Mismatches 49; Indels 39; Gaps 7;
XX
3 VQPNFNSVTLPGRLAGLALPRLPAHYQFLDDGVRAHLSLTERGPHSDSCGTLTLHR 42
192 VENGDFNWIIPGKITLSPGPHNESRENGPYHAPDYFYDFRENKYSTIVLAKKNYA 251

```

OY 43 ---TERGPPHSDSCPGTLRLRIPDCPPAPDIDRFVQIVDEANAGAVGHCALGF 99
 DB 252 SKFTKAGFDHVD-----LFTIDGSTPDEIMLKFIKVVDD--NRKG-GVAHVCKRAGL 299
 OY 100 GRGTMLACTIVKRGGLAGDAIAEIRLRPGST----ETYEOKAVF 143
 DB 300 GRGTGLIACMMKMEYGLTAGCMGLHYCRGYSVIGPQPYLIEKQKF 347

RESULT 11

AA67453
 ID AAG67453 standard; Protein: 447 AA.

AC AAG67453;

DT 26-NOV-2001 (first entry)

DE Amino acid sequence of a human polypeptide.

KW Human; protein kinase; protein phosphatase; signal transduction;

OS Homo sapiens.

PN WO200109345-A1.

PD 08-FEB-2001.

PF 28-JUL-2000; 2000MO-JP05060.

PR 29-JUL-1999; 99JP-0248036.

PR 18-OCT-1999; 99US-0159590.

PR 11-JAN-2000; 2000JP-0118776.

PR 17-FEB-2000; 2000US-0183322.

PR 02-MAY-2000; 2000JP-0183767.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;

PI Senoo C, Nezu J;

DR WPI; 2001-564736/63.

XX New genes encoding protein kinase and protein phosphatase, useful for

PT identifying modulators which can be used to treat human or animal

PT disorders associated with the expression or function of these enzymes -

PS Example 4; Page 291-294; 336pp; Japanese.

CC The specification describes human protein kinase/protein phosphatases.

CC in cells. The kinase phosphatases are connected with intracellular

CC signaling pathways. Antisense oligonucleotides and compounds

CC identified by screening (agonists or antagonists) can be used to

CC treat human or animal disorders associated with the expression

CC or function of the protein. In addition, the polypeptides may be used

CC as target molecules for drug development. The present sequence

CC represents a polypeptide, used in the course of the invention.

XX SQ Sequence 447 AA;

Query Match 21.4%; Score 170.5; DB 22; Length 447;

Best Local Similarity 29.4%; Pred. No. 3.9e-11; Indels 31; Gaps 7;

Matches 50; Conservative 27; Mismatches 62;

OY 7 NFSWVLPGRLAGLALP---RLPAHY-----QFLDLGVRLHLSLTER---GPPHS 50

DB 185 DLNMIIPDRIRACGPHSRARLESQYHSHSPETIYQFKHNNTTITRLKKRYDAKRF 244

OY 51 DSCPGTLRLRLRIPDCPPAPDIDRFVQIVDEANAGAVGHCALGFRTGTMLACYL 110

DB 245 DA--GFDHDLFPADGSTPDIAIVKRFIDICE--NAEG-AIAVHCKRAGLRTGLIACYI 299

DB 245 DA--GFDHDLFPADGSTPDIAIVKRFIDICE--NAEG-AIAVHCKRAGLRTGLIACYI 299
 OY 111 VKERGLAAGDAIAEIRLRPGSIETYEOKAVFO-----FYQRTK 150
 DB 300 MKHYRMTAAETIAVWRICRPLVIGPQOFLVMKRTSLMLBGDFRQRLK 349

RESULT 12

AA67632
 ID AAG67632 standard; Protein: 447 AA.

AC AAG67632;

DT 26-NOV-2001 (first entry)

DE Amino acid sequence of a human protein.

KW Human; protein kinase; protein phosphatase; signal transduction.

OS Homo sapiens.

PN WO200109316-A1.

PD 08-FEB-2001.

PF 28-JUL-2000; 2000MO-JP05061.

PR 29-JUL-1999; 99JP-0248036.

PR 18-OCT-1999; 99US-0159590.

PR 11-JAN-2000; 2000JP-0118776.

PR 17-FEB-2000; 2000US-0183322.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;

PI Senoo C, Nezu J;

DR WPI; 2001-570286/64.

XX New genes encoding proteins with protein kinase/protein phosphatase

PT activity, useful in the diagnosis and treatment of diseases -

PS Example 4; Page 191-194; 233pp; Japanese.

CC The specification describes human protein kinase/protein phosphatases.

CC It is expected that the protein kinase/protein phosphatase gene

CC participates in signal transduction in cells. The protein

CC kinase/protein phosphatase polypeptides and polynucleotides are

CC useful for developing diagnostics and treatment agents for human

CC and animal diseases. The protein kinase/protein phosphatase polypeptides

CC are useful as target molecules in designing novel drugs. The protein

CC kinase/protein phosphatase polynucleotides are useful as a source of

CC probes and primers, which may be used to isolate homologous sequences.

CC The present sequence represents a human protein, which is used in the

XX SQ Sequence 447 AA;

Query Match 21.4%; Score 170.5; DB 22; Length 447;

Best Local Similarity 29.4%; Pred. No. 3.9e-11; Indels 31; Gaps 7;

Matches 50; Conservative 27; Mismatches 62;

OY 7 NFSWVLPGRLAGLALP---RLPAHY-----QFLDLGVRLHLSLTER---GPPHS 50

DB 185 DLNMIIPDRIRACGPHSRARLESQYHSHSPETIYQFKHNNTTITRLKKRYDAKRF 244

OY 51 DSCPGTLRLRLRIPDCPPAPDIDRFVQIVDEANAGAVGHCALGFRTGTMLACYL 110

DB 245 DA--GFDHDLFPADGSTPDIAIVKRFIDICE--NAEG-AIAVHCKRAGLRTGLIACYI 299

```

OY 111 VKERGLAAGDAIAEIRLRPGSIEFYOEKAVFO-----FYORFK 150
DB 300 MKHRTMRAETIAWVRICRPGELVGPQOQFLVMKOTSLMEGDFRQRLK 349

RESULT 13
AAU75362
ID AAU75362 standard; Protein: 578 AA.
AC AAU75362;
DT 09-APR-2002 (first entry)
XX

DE Human dual specificity phosphatase CDC14A deletion variant.
XX
KW Human; cell-cycle control; CDC14A; cancer; deletion variant;
KW prostate cancer; breast cancer; tumour; lymph node metastasis;
KW malignant mesothelioma; chromosome 1p21; dual specificity phosphatase;
KW gene therapy; protein replacement therapy.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT MISC-difference 417..418
XX /note="Wild-type RSDPTKGRHRAVSQPPRL substituted by RL"
XX
XX OS US6331614-B1.
XX PD 18-DEC-2001.
XX PE 22-DEC-1999; 99US-0468872.
XX PR 23-DEC-1998; 98US-113833P.
XX PA (MYRI-) MYRIAD GENETICS INC.
XX PI Wong AKC, Teng DHF, Tavtigian SV;
XX DR WPI; 2002-129551/17.
XX DR N-PSDB; ABK13395.
XX

PT Nucleic acid encoding mutated form of human dual-specificity
PT phosphatase CDC14A polypeptide, useful to diagnose and treat cancers
PS Example 1; Page -: 41pp; English.
XX
XX CC The invention relates to an isolated nucleic acid encoding a CDC14A
XX CC polypeptide (cell-cycle control protein 14A, a dual specificity
XX CC phosphatase), its complement or RNA molecule corresponding to it.
XX CC Also included are an expression vector comprising the nucleic acid
XX CC and a host cell transformed with the vector. The gene for CDC14A is
XX CC located on human chromosome 1p21. The nucleic acid and protein are useful
XX CC to diagnose and treat human cancers (e.g. breast cancer, prostate
XX CC cancer) and tumours (e.g. lymph node metastasis, malignant mesothelioma)
XX CC which have a mutation in the CDC14A gene, by gene therapy, protein
XX CC replacement therapy or protein mimetics. They can also be used to
XX CC screen for drugs to treat cancer. The present sequence represents a
XX CC deletion variant of CDC14A encoded by a cDNA isolated from a breast cell
XX CC line.
XX CC Note: The present sequence is not shown in the specification but was
XX CC created by the indexer using the CDC14A sequence appearing as AAU75361
XX CC and the information in table 4.
XX
SQ Sequence 578 AA;

Query Match 20.1%; Score 160.5; DB 23; Length 578;
Best Local Similarity 29.9%; Pred. No. 8.1e-10; Indels 35; Gaps 5;
Matches 46; Conservative 22; Mismatches 51; Indels 35; Gaps 5;

OY 3 VQPPNFSWVLPGRLLA-----GLALPRLPAHYQFLDLGVNHLVSL----- 42
DB 175 VENGDFNMTIVGKFLAASGPHPKSKIENGYPPLAEPAYFFPKKHNTVAAYVRLKKKIYEA 234
OY 175 VENGDFNMTIVGKFLAASGPHPKSKIENGYPPLAEPAYFFPKKHNTVAAYVRLKKKIYEA 234

```

```

OY 43 ---TERGPPHSDSCGLTLHRLRIPDFCPAPDQIDRFVQYVDENANGAEGVHCALGF 99
DB 235 KRFTDAGEHND-----LFFIDGSTPSPDNIVRRLNICE--NTEG-AIAVHCKAGL 282

OY 100 GRTGTMLACYLVKERGLAAGDAIAEIRLRPGSI 133
DB 283 GRTGTMLACYLVKRTFTHAETIAWVRICRPGSI 316

RESULT 14
AAU75361
ID AAU75361 standard; Protein: 594 AA.
AC AAU75361;
DT 09-APR-2002 (first entry)
XX

DE Human dual specificity phosphatase CDC14A.
XX
KW Human; cell-cycle control; CDC14A; cancer;
KW prostate cancer; breast cancer; tumour; lymph node metastasis;
KW malignant mesothelioma; chromosome 1p21; dual specificity phosphatase;
KW gene therapy; protein replacement therapy.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT MISC-difference 417..418
XX /note="Wild-type RSDPTKGRHRAVSQPPRL substituted by RL"
XX
XX OS US6331614-B1.
XX PD 18-DEC-2001.
XX PE 22-DEC-1999; 99US-0468872.
XX PR 23-DEC-1998; 98US-113833P.
XX PA (MYRI-) MYRIAD GENETICS INC.
XX PI Wong AKC, Teng DHF, Tavtigian SV;
XX DR WPI; 2002-129551/17.
XX DR N-PSDB; ABK13392.
XX

PT Nucleic acid encoding mutated form of human dual-specificity
PT phosphatase CDC14A polypeptide, useful to diagnose and treat cancers
PS Claim 1; Column 58-62; 41pp; English.
XX
XX CC The invention relates to an isolated nucleic acid encoding a CDC14A
XX CC polypeptide (cell-cycle control protein 14A, a dual specificity
XX CC phosphatase), its complement or RNA molecule corresponding to it.
XX CC Also included are an expression vector comprising the nucleic acid
XX CC and a host cell transformed with the vector. The gene for CDC14A is
XX CC located on human chromosome 1p21. The nucleic acid and protein are useful
XX CC to diagnose and treat human cancers (e.g. breast cancer, prostate
XX CC cancer) and tumours (e.g. lymph node metastasis, malignant mesothelioma)
XX CC which have a mutation in the CDC14A gene, by gene therapy, protein
XX CC replacement therapy or protein mimetics. They can also be used to
XX CC screen for drugs to treat cancer. The present sequence represents CDC14A.
XX
SQ Sequence 594 AA;

Query Match 20.1%; Score 160.5; DB 23; Length 594;
Best Local Similarity 29.9%; Pred. No. 8.4e-10; Indels 35; Gaps 5;
Matches 46; Conservative 22; Mismatches 51; Indels 35; Gaps 5;

OY 3 VQPPNFSWVLPGRLLA-----GLALPRLPAHYQFLDLGVNHLVSL----- 42
DB 175 VENGDFNMTIVGKFLAASGPHPKSKIENGYPPLAEPAYFFPKKHNTVAAYVRLKKKIYEA 234
OY 43 ---TERGPPHSDSCGLTLHRLRIPDFCPAPDQIDRFVQYVDENANGAEGVHCALGF 99
DB 235 KRFTDAGEHND-----LFFIDGSTPSPDNIVRRLNICE--NTEG-AIAVHCKAGL 282
OY 100 GRTGTMLACYLVKERGLAAGDAIAEIRLRPGSI 133

```

Db 283 GRGTGLIACYMKHYRFTHAETIAMIRICRPGSI 316

RESULT 15

AAU75363

ID AAU75363 standard; Protein: 594 AA.

AC AAU75363;

DT 09-APR-2002 (first entry)

DE Human dual specificity phosphatase CDC14A G571R variant.

XX Human; cell-cycle control; CDC14A; cancer; G571R;

KW prostate cancer; breast cancer; tumour; lymph node metastasis; malignant mesothelioma; chromosome 1p21; dual specificity phosphatase; gene therapy; protein replacement therapy.

KM

XX Homo sapiens.

FH Key Location/Qualifiers

FT Misc-difference 571 /note="Wild-type Gly substituted by Arg"

FN US6331614-B1.

PD 18-DEC-2001.

PF 22-DEC-1999; 990S-0468872.

PR 23-DEC-1998; 980S-113833P.

PA (MYRI-) MYRIAD GENETICS INC.

PI Wong AKC, Teng DHF, Tavtigian SV;

DR WPI; 2002-129551/17.

DR N-PSDB; ABK13397.

XX Nucleic acid encoding mutated form of human dual-specificity phosphatase CDC14A polypeptide, useful to diagnose and treat cancers -

PS Example 1; Page -; 41pp; English.

XX The invention relates to an isolated nucleic acid encoding a CDC14A polypeptide (cell-cycle control protein 14A, a dual specificity

XX phosphatase), its complement or RNA molecule corresponding to it.

XX Also included are an expression vector comprising the nucleic acid

XX and a host cell transformed with the vector. The gene for CDC14A is

XX located on human chromosome 1p21. The nucleic acid and protein are useful

XX to diagnose and treat human cancers (e.g. breast cancer, prostate

XX cancer) and tumours (e.g. lymph node metastasis, malignant mesothelioma)

XX which have a mutation in the CDC14A gene, by gene therapy, protein

XX replacement therapy or protein mimetics. They can also be used to

XX screen for drugs to treat cancer. The present sequence represents a

XX variant of CDC14A.

XX Note: The present sequence is not shown in the specification but was

XX created by the indexer using the CDC14A sequence appearing as AAU75361

XX and the information in table 4.

SQ Sequence 594 AA;

Query Match 20.1%; Score 160.5; DB 23; Length 594;

Best Local Similarity 29.9%; Pred. No. 8.4e-10; Matches 46; Conservative 22; Mismatches 51; Indels 35; Gaps 5;

OY 3 VOPPNFSWVLPGRLA-----GLALPRLPAHYOFLDLGVRLVSL-----42

DB 175 VENGDPMNMTIVPGKFLAFSGHPKSKLENGYPLAPAEAFYFFKKNYTAAYVRLNKITYEA 234

OY 43 ----TENGPHSDSCPGILTLRLRIIPDFCPAPQIDRFVQIVDEANRGEAVGVHCALGF 99

Db 235 KRFTDAGFEHYD-----LFFIDGSTPSDNIVRRLINICE--NTEG-AIAYHCKAGL 282

OY 100 GRGTGLIACYMKHYRFTHAETIAMIRICRPGSI 133

Db 283 GRGTGLIACYMKHYRFTHAETIAMIRICRPGSI 316

Search completed: November 4, 2002, 12:46:16
Job time : 69 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Comugen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 2002, 12:44:52 ; Search time 26 seconds
(without alignments)
169.748 Million cell updates/sec

Title: US-09-619-380-2

Perfect score: 798
Sequence: 1 MGVPNFSWVLPGRILA.....GSIETVEOKAVFOFYORTK 150

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRMUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	160.5	20.1	594	US-09-468-872-2	Sequence 2, Appl
2	128.5	16.1	196	US-09-752-165-95	Sequence 95, Appl
3	117.5	14.7	173	US-09-704-139-4	Sequence 4, Appl
4	115	14.4	195	US-09-752-165-93	Sequence 93, Appl
5	115	14.4	195	US-09-752-165-96	Sequence 96, Appl
6	115	14.4	597	US-09-134-218-4	Sequence 4, Appl
7	115	14.4	597	US-09-134-218-5	Sequence 5, Appl
8	114	14.3	168	US-09-188-579-85	Sequence 85, Appl
9	114	14.3	168	US-09-315-444-85	Sequence 85, Appl
10	114	14.3	168	US-09-721-362-85	Sequence 85, Appl
11	114	14.3	168	US-09-752-165-99	Sequence 99, Appl
12	113	14.2	197	US-09-752-165-94	Sequence 94, Appl
13	112	14.0	173	US-09-315-444-87	Sequence 87, Appl
14	112	14.0	173	US-09-188-579-87	Sequence 87, Appl
15	112	14.0	173	US-09-721-362-87	Sequence 87, Appl
16	111.5	14.0	353	US-09-013-881-3	Sequence 3, Appl
17	111	13.9	172	US-09-704-139-5	Sequence 5, Appl
18	110	13.8	201	US-08-933-750C-21	Sequence 21, Appl
19	110	13.8	201	US-09-234-613-21	Sequence 21, Appl
20	107	13.4	176	US-09-704-139-2	Sequence 2, Appl
21	105.5	13.2	453	US-08-951-260A-7	Sequence 7, Appl
22	104	13.0	263	US-08-685-992-13	Sequence 13, Appl
23	104	13.0	263	US-09-144-925-13	Sequence 13, Appl
24	104	13.0	595	US-08-202-389-6	Sequence 6, Appl
25	104	13.0	631	US-08-202-389-8	Sequence 8, Appl
26	103.5	13.0	453	US-08-821-278A-2	Sequence 2, Appl
27	102.5	12.8	173	US-09-164-193-2	Sequence 2, Appl

28	102.5	12.8	173	US-09-221-448A-2	Sequence 2, Appl
29	102	12.8	175	US-09-188-579-86	Sequence 86, Appl
30	102	12.8	175	US-09-315-444-86	Sequence 86, Appl
31	102	12.8	175	US-09-721-362-86	Sequence 86, Appl
32	102	12.8	513	US-08-202-389-2	Sequence 2, Appl
33	101.5	12.7	369	US-08-596-291-4	Sequence 4, Appl
34	101.5	12.7	369	US-09-100-804-4	Sequence 4, Appl
35	101.5	12.7	369	US-09-100-804-4	Sequence 4, Appl
36	98.5	12.3	699	US-08-348-006B-7	Sequence 7, Appl
37	98.5	12.3	699	US-08-800-825R-7	Sequence 7, Appl
38	98.5	12.3	699	US-09-158-657-7	Sequence 7, Appl
39	98	12.3	249	US-08-685-992-28	Sequence 28, Appl
40	98	12.3	249	US-09-144-925-28	Sequence 28, Appl
41	98	12.3	360	US-09-020-743-2	Sequence 2, Appl
42	96.5	12.1	150	US-09-164-193-23	Sequence 23, Appl
43	96.5	12.1	150	US-09-221-448A-23	Sequence 23, Appl
44	95.5	12.0	528	US-08-987-691A-4	Sequence 4, Appl
45	94	11.8	278	US-08-821-278A-18	Sequence 18, Appl

ALIGNMENTS

```
RESULT 1
US-09-468-872-2
; Sequence 2, Application US/09468872
; Patent No. 6331614
;
GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Teng, David H.F.
; APPLICANT: Tavligian, Sean V.
; TITLE OF INVENTION: Human CDC14A Gene
; FILE REFERENCE: CDC14A Gene
; CURRENT APPLICATION NUMBER: US/09/468,872
; CURRENT FILING DATE: 1999-12-22
; EARLIER APPLICATION NUMBER: US 60/113,833
; EARLIER FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-468-872-2

Query Match      20.1%; Score 160.5; DB 4; Length 594;
Best Local Similarity 29.9%; Pred. No. 2.7e-11;
Matches 46; Conservative 22; Mismatches 51; Indels 35; Gaps 5;

QY      3  VQPPNFSWVLPGRILA-----GLAUPRLRAHQFLDLGVRHLYSL----- 42
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      175  VENGDMNVIQGFLEAFSGDPHPKSKLENGYPLHAPRAVPEYFKKHVTAVRLNKKIYEA 234
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      43  ---TERGPPHSDSCPGILTLRLRIPDFCPAPDQIDRFVOYDEANARGEAVGHCALGF 99
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      235  KRRTDAGFEHYD-----LEFTGSPSPDNIIVRFNLICE--NTEG-ALNVHCKAGL 282
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      100  GRGTGLACIYKVERGLAAGDAIAETRLRLPGSI 133
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      283  GRGTGLACIYKVRFTFAETIRIATIRICRPGSI 316
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
US-09-752-165-95
; Sequence 95, Application US/09752165
; Patent No. 6451583
;
GENERAL INFORMATION:
; APPLICANT: Shuman, Stewart
; APPLICANT: Ho, C. Kiong
; TITLE OF INVENTION: Pharmacological Targeting Of mRNA Cap Formation
; TITLE OF INVENTION: For treatment of Parasitic Infections
; FILE REFERENCE: D6388
; CURRENT APPLICATION NUMBER: US/09/752,165
```



```

; TITLE OF INVENTION: mRNA CAPPING ENZYMES AND USES THEREOF
; FILE REFERENCE: 601-1-079 ss
; CURRENT APPLICATION NUMBER: US/09/134,218A
; CURRENT FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-134-218-4

Query Match
Best local Similarity 14.4%; Score 115; DB 4; Length 597;
Matches 24; Conservative 18; Mismatches 36; Indels 2; Gaps 1;

QY 69 PARDQIDRFVQIVDEANAR--EAVGVHCAAGFGRTGTMACYIKERGLAAGDAIEIR 126
DB 98 PTENTETFRICERFERNRPPELLIGVHCTHGNRTGFLICAPLVERKMDWSIEAAVATFA 157
127 RLRRGSIETYEOKAVFOFY 146
DB 158 QARPPGIYKGDYKELFRRY 177

RESULT 7
US-09-134-218-5
; Sequence 5, Application US/09134218A
; Patent No. 6312926
; GENERAL INFORMATION:
; APPLICANT: Shatkin, Aaron J.
; APPLICANT: Pillutla, Renuka
; APPLICANT: Reinberg, Danny
; APPLICANT: Yu, Zheng
; APPLICANT: Moldonado, Edio
; TITLE OF INVENTION: mRNA CAPPING ENZYMES AND USES THEREOF
; FILE REFERENCE: 601-1-079 ss
; CURRENT APPLICATION NUMBER: US/09/134,218A
; CURRENT FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-134-218-5

Query Match
Best local Similarity 14.4%; Score 115; DB 4; Length 597;
Matches 24; Conservative 18; Mismatches 36; Indels 2; Gaps 1;

QY 69 PARDQIDRFVQIVDEANAR--EAVGVHCAAGFGRTGTMACYIKERGLAAGDAIEIR 126
DB 98 PTENTETFRICERFERNRPPELLIGVHCTHGNRTGFLICAPLVERKMDWSIEAAVATFA 157
127 RLRRGSIETYEOKAVFOFY 146
DB 158 QARPPGIYKGDYKELFRRY 177

RESULT 8
US-09-188-579-85
; Sequence 85, Application US/09188579B
; Patent No. 6107040
; GENERAL INFORMATION:
; APPLICANT: Shuman, Stewart
; TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
; FILE REFERENCE: D6185
; CURRENT APPLICATION NUMBER: US/09/188,579B
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 114
; SEQ ID NO 85
; LENGTH: 168
```

```

; TYPE: PRT
; ORGANISM: baculovirus
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of RNA baculovirus phosphatase BVP.
US-09-188-579-85

Query Match
Best local Similarity 14.3%; Score 114; DB 3; Length 168;
Matches 29; Conservative 13; Mismatches 43; Indels 4; Gaps 2;

QY 55 GLTLRLRIPDFCPAPDQIDRFVQIVDEANAR--GEAVGVHCAAGFGRTGTMACYIK 112
DB 77 GLYKKIQVGGQLPPEISVQEFIDYVKETKCPGMLVGVHCTHGINRTGYMCRYLHM 136
113 ERGLAAGDAIEIRRLRRPSIE--TYROE 139
DB 137 TLGIAPOEAIIDREFKARGHKIERONYOD 165

RESULT 9
US-09-315-444-85
; Sequence 85, Application US/09315444A
; Patent No. 6232070
; GENERAL INFORMATION:
; APPLICANT: Shuman, Stewart
; TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
; FILE REFERENCE: D6185CIP
; CURRENT APPLICATION NUMBER: US/09/315,444A
; CURRENT FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 09/188,579
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 85
; LENGTH: 168
; TYPE: PRT
; ORGANISM: baculovirus
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of RNA baculovirus phosphatase BVP.
US-09-315-444-85

Query Match
Best local Similarity 14.3%; Score 114; DB 4; Length 168;
Matches 29; Conservative 13; Mismatches 43; Indels 4; Gaps 2;

QY 55 GLTLRLRIPDFCPAPDQIDRFVQIVDEANAR--GEAVGVHCAAGFGRTGTMACYIK 112
DB 77 GLYKKIQVGGQLPPEISVQEFIDYVKETKCPGMLVGVHCTHGINRTGYMCRYLHM 136
113 ERGLAAGDAIEIRRLRRPSIE--TYROE 139
DB 137 TLGIAPOEAIIDREFKARGHKIERONYOD 165

RESULT 10
US-09-721-362-85
; Sequence 85, Application US/09721362
; Patent No. 6420163
; GENERAL INFORMATION:
; APPLICANT: Shuman, Stewart
; TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
; FILE REFERENCE: D6185CIP/D
; CURRENT APPLICATION NUMBER: US/09/721,362
; CURRENT FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 09/315,444
; PRIOR FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 85
; LENGTH: 168
; TYPE: PRT
; ORGANISM: baculovirus
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of RNA baculovirus phosphatase BVP.
US-09-721-362-85
```

Query Match 14.3%; Score 114; DB 4; Length 168;
Best Local Similarity 32.6%; Pred. No. 2,7e-06;
Matches 29; Conservative 13; Mismatches 43; Indels 4; Gaps 2;
QY 55 GLTHRLRIPDFCPADQIDRFVQIYDEANAR--GEAVGVHCALGFGRTGTMACYLTK 112
DB 77 GLYKKIQVGGQTPPESSIVGEFIDYKFEKPCGMLGVGCHTGINRTGYMCRTLMH 136
QY 113 ERGLAAGDAIAEIRLRPGSIE--TYBOE 139
DB 137 TLGIAPQEAIDRFKARKHRIEKNYOD 165
RESULT 11
US-09-752-165-99
Sequence 99; Application US/09/752165
Patent No. 6451583
GENERAL INFORMATION:
APPLICANT: Shuman, Stewart
APPLICANT: Ho, C. Kiong
TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
FILE REFERENCE: D6388
CURRENT FILING DATE: 2000-12-29
CURRENT APPLICATION NUMBER: US/09/752,165
NUMBER OF SEQ ID NOS: 104
SEQ ID NO 99
LENGTH: 168
TYPE: PRT
ORGANISM: Unknown
FEATURE:
NAME/KEY: DOMAIN
OTHER INFORMATION: Baculovirus RNA-specific 5' phosphatase BVP
US-09-752-165-99
Query Match 14.3%; Score 114; DB 4; Length 168;
Best Local Similarity 32.6%; Pred. No. 2,7e-06;
Matches 29; Conservative 13; Mismatches 43; Indels 4; Gaps 2;
QY 55 GLTHRLRIPDFCPADQIDRFVQIYDEANAR--GEAVGVHCALGFGRTGTMACYLTK 112
DB 77 GLYKKIQVGGQTPPESSIVGEFIDYKFEKPCGMLGVGCHTGINRTGYMCRTLMH 136
QY 113 ERGLAAGDAIAEIRLRPGSIE--TYBOE 139
DB 137 TLGIAPQEAIDRFKARKHRIEKNYOD 165
RESULT 12
US-09-752-165-94
Sequence 94; Application US/09/752165
Patent No. 6451583
GENERAL INFORMATION:
APPLICANT: Shuman, Stewart
APPLICANT: Ho, C. Kiong
TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
FILE REFERENCE: D6388
CURRENT APPLICATION NUMBER: US/09/752,165
CURRENT FILING DATE: 2000-12-29
NUMBER OF SEQ ID NOS: 104
SEQ ID NO 94
LENGTH: 197
TYPE: PRT
ORGANISM: C elegans
FEATURE:
NAME/KEY: DOMAIN
OTHER INFORMATION: RNA triphosphatase domain of C elegans RNA capping enzyme
US-09-752-165-94
Query Match 14.2%; Score 113; DB 4; Length 197;
Best Local Similarity 36.0%; Pred. No. 4,4e-06;

Matches 31; Conservative 12; Mismatches 37; Indels 6; Gaps 3;
QY 69 PAPQIDRFVQIYDEANAR--GEAVGVHCALGFGRTGTMACYL--VKERGLAAGDAIAE 124
DB 95 PTOEDTNEFKLVQEFHKKYPDRVGVGCHTGFNRTGFLIAAYLFOVEEYGLDA--AIGE 152
QY 125 IRLRPGSIEEYEOEKAVFOFYORTK 150
DB 153 FAENRKGIVKQDYIDLFARNDPTE 178
RESULT 13
US-09-188-579-87
Sequence 87; Application US/09188579B
Patent No. 6107040
GENERAL INFORMATION:
APPLICANT: Shuman, Stewart
TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
FILE REFERENCE: D6185
CURRENT APPLICATION NUMBER: US/09/188,579B
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 114
SEQ ID NO 87
LENGTH: 173
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
OTHER INFORMATION: Amino acid sequence of N-terminal RNA triphosphatase
US-09-188-579-87
Query Match 14.0%; Score 112; DB 3; Length 173;
Best Local Similarity 40.0%; Pred. No. 4,9e-06;
Matches 30; Conservative 11; Mismatches 26; Indels 8; Gaps 4;
QY 69 PAPQIDRFVQIYDEANAR--GEAVGVHCALGFGRTGTMACYL--VKERGLAAGDAIAE 124
DB 96 PTOEDTNEFKLVQEFHKKYPDRVGVGCHTGFNRTGFLIAAYLFOVEEYGLDA--AIGE 153
QY 125 IRLRPGSIEEYEOE 139
DB 154 FAENRKGIT--YKOD 166
RESULT 14
US-09-315-444-87
Sequence 87; Application US/09315444A
Patent No. 6232070
GENERAL INFORMATION:
APPLICANT: Shuman, Stewart
TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
FILE REFERENCE: D6185CJP
CURRENT APPLICATION NUMBER: US/09/315,444A
CURRENT FILING DATE: 1999-05-20
PRIOR APPLICATION NUMBER: US 09/188,579
PRIOR FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 116
SEQ ID NO 87
LENGTH: 173
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
OTHER INFORMATION: Amino acid sequence of N-terminal RNA triphosphatase
US-09-315-444-87

Query Match 14.0%; Score 112; DB 4; Length 173;
Best Local Similarity 40.0%; Pred. No. 4,9e-06;
Matches 30; Conservative 11; Mismatches 26; Indels 8; Gaps 4;
QY 69 PAPQIDRFVQIYDEANAR--GEAVGVHCALGFGRTGTMACYL--VKERGLAAGDAIAE 124
DB 96 PTOEDTNEFKLVQEFHKKYPDRVGVGCHTGFNRTGFLIAAYLFOVEEYGLDA--AIGE 153

```

OY      125  IRLRPGSIETYEQE 139
          |      | :|:
Db      154  FAENRÖKGI - YKÖD 166

```

RESULT 15
HE-00-731

```

US-09-721-362-87
: Sequence 87, Application US/09721362
: Patent NO. 6420163
: GENERAL INFORMATION:
: APPLICANT: Shuman, Stewart
: TITLE OF INVENTION: Pharmacological Targeting of mRNA Cap Formation
: FILE REFERENCE: D6183C/P/D
: CURRENT APPLICATION NUMBER: US/09/721,362
: CURRENT FILING DATE: 2000-11-22
: PRIOR APPLICATION NUMBER: US 09/315,444
: PRIOR FILING DATE: 1999-05-20
: NUMBER OF SEQ ID NOS: 116
: SEQ ID NO 87
: LENGTH: 173
: TYPE: PRT
: ORGANISM: Candida albicans
: FEATURE:
: OTHER INFORMATION: Amino acid sequence of N-terminal RNA triphosphatase
: OTHER INFORMATION: domain of the capping enzyme.
US-09-721-362-87

```

Query Match 14.0%; Score 112; DB 4; Length 173;
Best Local Similarity 40.0%; Pred. No. 4.9e-06;
Matches 30; Conservative 11; Mismatches 26; Indels 8; Gaps 4;

Qy	69	PAPQIDFEVQIVDEANAR--GEAVGHCACAFGRTGTMACY--VKESGIAAGDIAE	124
		: : : : : : : : : : : : : : : :	
	96	PTQEDTDMFNLVQGEFHKKRYDRVVGHCSTGFNRTGFLAAYLFLQYBEVGLDA--YIGE	133
		: : : : : : : : : : : : : : : :	
Qy	125	IRRLRPGSIEIYEQD	139
		: : : : : : : : : : : : : : : :	
Db	154	FAENRQKGI--YKQD	166

Search completed: November 4, 2002, 12:50:38
Job time : 28 secs

RESULT 2
US-09-815-419-4

```

: Sequence 4, Application US/09815419
: Patent No. US20020065406A1
: GENERAL INFORMATION:
: APPLICANT: Meyers, Rachel A.
: TITLE OF INVENTION: 18221, A NOVEL DUAL SPECIFICITY
: TITLE OF INVENTION: PHOSPHATASE AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: 10448-030001
: CURRENT APPLICATION NUMBER: US/09/815,419
: CURRENT FILING DATE: 2001-03-22
: PRIOR APPLICATION NUMBER: US 60/191,858
: PRIOR FILING DATE: 2000-03-24
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 173
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: consensus sequence
US-09-815-419-4

```

```

Query Match          14.7%; Score 117.5; DB 10; Length 173;
Best Local Similarity 25.7%; Pred. No. 1.1e-05;
Matches   37; Conservative    19; Mismatches   51; Indels     37; Gaps      4;

QY       9 SWVLPGRLAGLALPRPAHYOFLDLGVRHLVSITRGP-----PHSDSC 53
           |::|              ||:||::|||         :
DB        3 SEIPLHLYGSYSTASANTALLKLGITGVINVTVEVPNPFDLKNDRIHTAYNISKN 62
           |::|             |::|               |
QY       54 PGTTTLRLRPDPCPPAPDOI-----DRFQIVDEANAREANGVCAL 97
           |::|            |::|                |::|
DB        63 SGGTY--LOIPN---VDHIYYHANHEHKISKRYEADAVFDIDAAOKKGKVLVHCQA 116
           |::|            |::|                 |::|
QY       98 GFRTGTMLACYVKERGLAAGDA 121
           |::|            |::|                 |::|
DB       117 GISRSATLIATIAMKTIRNLISLENA 140
           |::|            |::|                 |::|


RESULT 3
US-09-964-277-7
Sequence 7, Application US/09964277
Patent No., US20020137170A1
GENERAL INFORMATION:
APPLICANT: Luche, Ralf M.
APPICANT: Wei, Bo
TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125_434
CURRENT APPLICATION NUMBER: US/09/964,277
CURRENT FILING DATE: 2001-09-25
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 155
TYPE: PRT
ORGANISM: Homo sapiens
US-09-964-277-7


Query Match          14.3%; Score 114; DB 10; Length 155;
Best Local Similarity 26.4%; Pred. No. 1.2e-05;
Matches   28; Conservative    20; Mismatches   42; Indels     16; Gaps      3;

QY       50 SDSCPQLTIL----HRLRIP---DFCCPAPQDIRFYQIVDEANARGAVGHCAIGFGRT 102
           |::|            |::|               |::|
DB        46 SVMTCRAPDPFPESHFLARVYNDSFCCEKLPMWLDKSVDIFEKAKASNCGVLVHCIAGISR 105
           |::|            |::|               |::|
QY       103 GTMALACTYVKERGLAAGDAIAETIRLRP-----GSTIEYEQE 139
           |::|            |::|               |::|
DB        106 AATAIAYIMKRMDMSIDEAYRFVFKEKKRPITSPTNFNIQGOLDYEKK 151
           |::|            |::|               |::|

```

```

: Sequence 21 Application US/09964277
: Patent No. US20020137170A1
: GENERAL INFORMATION:
: APPLICANT: LucHe, Ralf M.
: APPLICANT: Wei, Bo
: TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
: FILE REFERENCE: 200125.434
: CURRENT APPLICATION NUMBER: US/09/964,277
: CURRENT FILING DATE: 2001-09-25
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 21
: LENGTH: 517
: TYPE: PRM
: ORGANISM: Homo sapiens
: US-09-964-277-21

```

```

QY      50  SDSCBGLTL---HRLRIP---DCCPPADIDRFVQIVDEANARGAENVGHCALGFGR 102
          ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
Db      44  SNTCPKPDPIFESHFLRFVNDSEFCETILPWLKDSVDFIERAKSNGCVLHCLAGISRS 103
QY      103  GTMLACVLYKERGLAAGDAIAEIRLRP-----GSIETVBOE 139
          ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
Db      104  ATIAIATIMKRMDSLDEAIVRFVKEKRRTISPNFNFLGQLLDYERK 149

RESULT 5
US-09-816-494-2
; Sequence 2, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-816-494-2

Query Match      14.3%; Score 114; DB 10; Length 665;
Best Local Similarity 26.4%; Pred. No. 0.0001;
Matches 28; Conservative 20; Mismatches 42; Indels 16; Gaps 3;

QY      50  SDSCBGLTL---HRLRIP---DCCPPADIDRFVQIVDEANARGAENVGHCALGFGR 102
          ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
Db      192  SNTCPKPDPIFESHFLRFVNDSEFCETILPWLKDSVDFIERAKSNGCVLHCLAGISRS 251
QY      103  GTMLACVLYKERGLAAGDAIAEIRLRP-----GSIETVBOE 139
          ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
Db      252  ATIAIATIMKRMDSLDEAIVRFVKEKRRTISPNFNFLGQLLDYERK 297

RESULT 6
US-09-964-277-2
; Sequence 2, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434

```

```
;; CURRENT APPLICATION NUMBER: US/09/964,277
;; CURRENT FILING DATE: 2001-09-25
;; NUMBER OF SEQ ID NOS: 22
;; SOFTWARE: FastSeq for Windows Version 4.0.
;; SEQ ID NO 2
;; LENGTH: 665
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-964-277-2
```

```
Query Match
Best Local Similarity 26.4%; Pred. No. 0.00014;
Matches 28; Conservative 20; Mismatches 42; Indels 16; Gaps 3;
```

```
QY 50 SDSCPGITL-----RLRIRP---DECPAPQIDRFQIVDEANARGAVGHCALGFGRT 102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 192 SMCPRPDPFIFESHFLRPVNDSCFCEILPWLKSVDFIKAKRNSCGLVHCLAGISRS 251
```

```
QY 103 GTMLACYLVKERGLAGDAIAEIRLRP-----GSIETYEQ 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 252 ATIAIAYIMKRMDSIDEAVRFVKEKRPTISPENFGLQLEDER 297
```

```
RESULT 7
US-09-775-925-26
; Sequence 26, Application US/09775925
; Patent No. US20010049358A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Lucche, Ralf M.
```

```
; TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
```

```
; FILE REFERENCE: 200125.420
```

```
; CURRENT APPLICATION NUMBER: US/09/775,925
```

```
; CURRENT FILING DATE: 2001-02-01
```

```
; NUMBER OF SEQ ID NOS: 33
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 26
```

```
; LENGTH: 170
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
```

```
US-09-775-925-26
```

```
Query Match
Best Local Similarity 23.9%; Pred. No. 2.8e-05;
Matches 37; Conservative 29; Mismatches 62; Indels 27; Gaps 5;
```

```
QY 11 VLPGRLAGLALP-----RLPAHYOPLDLGVRLVSLTERGPPH-----SDSCPG----- 55
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10 LTPMSISQCLPVPVGLTRILPHLYLSQKDVLNKMLQNGISYVLNNSCCKPDPFI 69
```

```
QY 56 LTLHRLRIP---DECPAPQIDRFQIVDEANARGAVGHCALGFGRTGTMACYLVK 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 CESRFMRVPINDNYCEKLLPWLKDSIEFDIKAKLSQCVYVHCLAGISRSATIAIAYIMK 129
```

```
QY 113 ERGLAGDAIAEIRLRP-----GSIETYEQ 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 TMGMSDDAYRFVKDRPSPSPNFNFGQLETER 164
```

```
RESULT 8
US-09-847-519A-11
; Sequence 11, Application US/09847519A.
; Patent No. US20020102693A1
; GENERAL INFORMATION:
```

```
;; SEQ ID NO 11
;; LENGTH: 170
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-847-519A-11

Query Match
Best Local Similarity 14.2%; Score 113.5; DB 10; Length 170;
Matches 37; Conservative 29; Mismatches 62; Indels 27; Gaps 5;
```

```
QY 11 VLPGRLAGLALP-----RLPAHYOPLDLGVRLVSLTERGPPH-----SDSCPG----- 55
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10 LTPMSISQCLPVPVGLTRILPHLYLSQKDVLNKMLQNGISYVLNNSCCKPDPFI 69
```

```
QY 56 LTLHRLRIP---DECPAPQIDRFQIVDEANARGAVGHCALGFGRTGTMACYLVK 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 CESRFMRVPINDNYCEKLLPWLKDSIEFDIKAKLSQCVYVHCLAGISRSATIAIAYIMK 129
```

```
QY 113 ERGLAGDAIAEIRLRP-----GSIETYEQ 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 TMGMSDDAYRFVKDRPSPSPNFNFGQLETER 164
```

```
RESULT 9
US-09-964-277-6
; Sequence 6, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Lucche, Ralf M.
```

```
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
```

```
; FILE REFERENCE: 200125.434
```

```
; CURRENT APPLICATION NUMBER: US/09/964,277
```

```
; CURRENT FILING DATE: 2001-09-25
```

```
; NUMBER OF SEQ ID NOS: 22
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 6
```

```
; LENGTH: 155
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
```

```
US-09-964-277-6
```

```
Query Match
Best Local Similarity 14.0%; Score 112; DB 10; Length 155;
Matches 29; Conservative 25; Mismatches 43; Indels 20; Gaps 4;
```

```
QY 42 LTERGPPH-----SDSCPG-----LTLHRLRIP---DECPAPQIDRFQIVDEANARGEA 90
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34 MTONGISYVLNNSCCKPDPFICESRMRVPINDNYCEKLLPWLKDSIEFDIKAKLSQ 93
```

```
QY 91 VGYHCALGFGRTGTMACYLVKERGLAGDAIAEIRLRP-----GSIETYEQ 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 VIVHCLAGISRSATIAIAYIMKTMGMSDDAYRFVKDRPSPSPNFNFGQLETER 150
```

```
RESULT 10
US-09-816-494-8
; Sequence 8, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 172
; TYPE: PRT
```


APPLICANT: FLANAGAN, PETER
TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
FILE REFERENCE: 038602/1180
CURRENT APPLICATION NUMBER: US/09/866,987
CURRENT FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 60/208,291
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 201
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: phosphatase polypeptide
US-09-866-987-7

Query Match 13.8%; Score 110; DB 10; Length 201;
Best Local Similarity 27.8%; Pred. No. 8.5e-05;
Matches 44; Conservative 25; Mismatches 61; Indels 28; Gaps 7;

QY 8 FSNVLR-----LAGLALRLPAHQFLDLGVRLVSLTERGPPH----- 49
DB 25 FRKVPGRARHWYHRIDPTVLG-ALPLRLARQVQDENRGVITMNEEYETRFCLNS 83
QY 50 SDSCPGTLRLRLRP--DFCP-PAPQIDRFVQIVDEANRGEAVGVCHALGFGRTGTM 106
DB 84 SQMKRLGVQLRLSTYDMGIRPLDMLQKGVDPALKYSGOCVVYHCKAGRSRATMV 143
QY 107 ACYLKERGLA---AGDAIEIR---RLRPGSIETYEQ 138
DB 144 AAYLIQVHKVSPREAVRAIAKIRSYIHIRGQDVLKE 181

RESULT 14
US-09-775-925-25
Sequence 25, Application US/09775925
Patent No. US20010049358A1
GENERAL INFORMATION:
APPLICANT: Lucbe, Ralf M.
APPLICANT: Wei, Bo
TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
FILE REFERENCE: 200125.420
CURRENT APPLICATION NUMBER: US/09/775,925
CURRENT FILING DATE: 2001-02-01
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 157
TYPE: PRT
ORGANISM: Homo sapiens
US-09-775-925-25

Query Match 13.4%; Score 107; DB 10; Length 157;
Best Local Similarity 25.5%; Pred. No. 0.00013;
Matches 36; Conservative 25; Mismatches 56; Indels 24; Gaps 4;

QY 2 GVOPPNFS-----WVLPGRLAGLALRLPAHYQFLDLGVRLVSLTERGPPHSDSCP 54
DB 1 GATPPVGLRASPVOILPMLYLGSA--RDSANLESIAKIGIRYILNVTPLNLPFFFEKNG 58
QY 55 GLTLHLRLRPD-----FCPPAPQIDRFVQIVDEANRGEAVGVCHALGFGRTGTM 106
DB 59 DFHYKQIPISDHWSONLSRFFPEA-----IEFIDALSQNGCVLVHCLAGVRSRYVT 111
QY 107 ACYLKERGLAAGDAIEIR 127
DB 112 VAYLMQKHLSLNDAYDVKR 132

RESULT 15

US-09-847-519A-10
Sequence 10, Application US/09847519A
Patent No. US20020102693A1
GENERAL INFORMATION:
APPLICANT: Lucbe, Ralf M.
APPLICANT: Wei, Bo
TITLE OF INVENTION: DSP-14 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.422
CURRENT APPLICATION NUMBER: US/09/847,519A
CURRENT FILING DATE: 2001-05-01
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 157
TYPE: PRT
ORGANISM: Homo sapiens
US-09-847-519A-10

Query Match 13.4%; Score 107; DB 10; Length 157;
Best Local Similarity 25.5%; Pred. No. 0.00013;
Matches 36; Conservative 25; Mismatches 56; Indels 24; Gaps 4;

QY 2 GVOPPNFS-----WVLPGRLAGLALRLPAHYQFLDLGVRLVSLTERGPPHSDSCP 54
DB 1 GATPPVGLRASPVOILPMLYLGSA--RDSANLESIAKIGIRYILNVTPLNLPFFFEKNG 58
QY 55 GLTLHLRLRPD-----FCPPAPQIDRFVQIVDEANRGEAVGVCHALGFGRTGTM 106
DB 59 DFHYKQIPISDHWSONLSRFFPEA-----IEFIDALSQNGCVLVHCLAGVRSRYVT 111
QY 107 ACYLKERGLAAGDAIEIR 127
DB 112 VAYLMQKHLSLNDAYDVKR 132

Search completed: November 4, 2002, 12:57:06
Job time : 21 secs

This Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 2002, 07:42:22 ; Search time 2785 seconds

(without alignments)
7388.037 Million cell updates/sec

Title: US-09-619-380-1

Perfect score: 707

Sequence: 1 tgaccgcgtctcctgtgcc.....aaaaaaaaaaaaaaaa 707

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```
Database :
1:  gb_ba:*
2:  gb_hg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hgt_hum:*
31: em_hgt_inv:*
32: em_hgt_other:*
33: em_hgt_mus:*
34: em_hgt_pln:*
35: em_hgt_rtd:*
36: em_hgt_mam:*
37: em_hgt_vrt:*
38: em_sy:*
39: em_hgtgo_hum:*
40: em_hgtgo_mus:*
41: em_hgtgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	707	100.0	707	6	AX078400
2	703.8	99.5	878	6	AX482482
3	703	99.4	839	6	AX099936
4	679.8	96.2	711	9	BC001140
5	673.6	95.3	746	6	AX398840
6	673.6	95.3	746	6	AX398843
7	673.6	95.3	746	6	AX482336
8	673.6	95.3	746	6	AX000449
9	671.6	95.0	709	6	AX398833
10	657.4	93.0	694	6	AX086042
11	583.8	82.6	602	6	AX398836
12	555.8	78.6	559	9	HUM2E12005
13	508.6	71.9	553	6	AX398833
14	486.2	68.8	433	6	AX398834
15	446	63.1	448	6	AX398839
16	429	60.7	453	6	AX398835
17	396.4	56.1	715	6	AX086004
18	391	55.3	13862	9	AL580560
19	379.8	53.7	397	6	AX398837
20	373	52.8	433	6	AX078411
21	336.4	47.6	399	6	AX398830
22	241.8	34.2	247	6	AX334867
23	229.6	32.5	176846	2	AC131177
24	80	11.3	80	6	AX482385
25	79	11.2	79	6	AX482382
26	71.8	10.2	35694	3	CIC0541
27	57.4	8.1	36583	1	SC5H1
28	56.8	8.0	148766	2	AC118895
29	56.2	7.9	35033	1	SC130A
30	55.2	7.8	1146	6	I92666
31	55.2	7.8	2544	1	SA017169
32	55.2	7.8	2728	6	I92670
33	54.6	7.7	125020	9	AF429315
34	54	7.6	134350	2	AC105734
35	54	7.6	159636	8	AC090485
36	52.8	7.5	40337	1	SC5A7
37	52.8	7.5	110000	2	AC116301-0
38	52.6	7.4	83582	2	AC121741
39	52.6	7.4	104627	2	AC117882
40	52.2	7.4	54750	2	AC009072
41	52.2	7.4	214791	9	AC007610
42	52	7.4	79419	2	AC128943
43	51.2	7.2	69165	2	AC023465
44	51	7.2	10961	1	AE001939
45	50.8	7.2	20021	1	AE004730

ALIGNMENTS

RESULT 1
AX078400
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 707)
Luche, R.M. and Wei, B.
TITLE
Dsp-11 dual-specificity map kinase phosphatase
JOURNAL
Patent: WO 0105983-A 1 25-JAN-2001;

FEATURES Ceptyr, Inc. (US)
Location/Qualifiers
1. 707
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 146 a 224 c 202 g 135 t
ORIGIN

Query Match 100.0%; Score 707; DB 6; Length 707;
Best Local Similarity 100.0%; Pred. No. 5.6e-117;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGACCCGGTGTCTGCTGCTTCCAGAGATGGGCGTGCACCCCACTTCCTCG 60
DB 1 TGACCCGGTGTCTGCTGCTTCCAGAGATGGGCGTGCACCCCACTTCCTCG 60
QY 61 GTGCTCCGGGCGGCTGGCGGAGTGGCGCTGCGCGGCTCCCGCCCACTACAGTTG 120
DB 61 GTGCTCCGGGCGGCTGGCGGAGTGGCGCTGCGCGGCTCCCGCCCACTACAGTTG 120
QY 121 CTGTTGACCTGGGCGTGGCGACCTGTGCTCCAGAGAGCGGCGGCGGCTCAGAC 180
DB 121 CTGTTGACCTGGGCGTGGCGACCTGTGCTCCAGAGAGCGGCGGCGGCTCAGAC 180
QY 181 GACAGCTCCCGGCGGCTGACCTGACCGGCTGCGGCTCCCGGCTTCGCGCGGCG 240
DB 181 GACAGCTCCCGGCGGCTGACCTGACCGGCTGCGGCTCCCGGCTTCGCGCGGCG 240
QY 241 CCGGACAGATGACCGGCTGTGCGAGATCGAGAGAGCGGCGGCGGAGAGGCT 300
DB 241 CCGGACAGATGACCGGCTGTGCGAGATCGAGAGAGCGGCGGCGGAGAGGCT 300
QY 301 GTGGAGTGCACCTGTGCTGCGGCTTTGGCGGCTGCGGCTGCGGCTGTTACCTG 360
DB 301 GTGGAGTGCACCTGTGCTGCGGCTTTGGCGGCTGCGGCTGCGGCTGTTACCTG 360
QY 361 GTGAAGAGCGGCGGCTTGGCTGCGAGAGATGCGATTCGTAATCCGAGACTACGACC 420
DB 361 GTGAAGAGCGGCGGCTTGGCTGCGAGAGATGCGATTCGTAATCCGAGACTACGACC 420
QY 421 GGCCTCATCAGAGACTATGAGAGAGAGAGAGAGAGTTCGATTCAGAGAGAGAGAA 480
DB 421 GGCCTCATCAGAGACTATGAGAGAGAGAGAGAGAGTTCGATTCAGAGAGAGAGAA 480
QY 481 TAAGGGGCTTAGTACCTTCTACAGGCGCTCACTCCCTTCCCATGTGTCATGGG 540
DB 481 TAAGGGGCTTAGTACCTTCTACAGGCGCTCACTCCCTTCCCATGTGTCATGGG 540
QY 541 GCCAGAGTGAAGGAGAGTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 600
DB 541 GCCAGAGTGAAGGAGAGTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 600
QY 601 CACTGAAGTACCCACCCCTGAGAGAGTCTGTATTAAGGAGAGGCTTGTACTGCTTT 660
DB 601 CACTGAAGTACCCACCCCTGAGAGAGTCTGTATTAAGGAGAGGCTTGTACTGCTTT 660
QY 661 GTGAATTAATGAGTTTACGACCAAAAAAAAAAAAAAAAAAAAAA 707
DB 661 GTGAATTAATGAGTTTACGACCAAAAAAAAAAAAAAAAAAAAAA 707

RESULT 2
AX482482 878 bp DNA linear PAT 16-AUG-2002
LOCUS Sequence 151 from Patent WO02057460.
DEFINITION AX482482
ACCESSION AX482482
VERSION AX482482.1 GI:22317003
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Todderud,C.G., Bol,D., Finger,J., Lee,L., Nelson,T., Schieven,G.,
Suchard,S., Banas,D., Bassolino,D., Feder,J., Krystek,S.,
Mcatee,P., Minter,G., Siemers,N., Jackson,D.G. and Ramnathan,C.
TITLE Polynucleotides encoding human phosphatases
JOURNAL Patent: WO 02057460-A 151 25-JUL-2002;
SQUIBB BRISTOL MYERS CO (US)
FEATURES Location/Qualifiers
1. 878
/organism="Homo sapiens"
/db_xref="taxon:9606"
89. 541
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD44455.1"
/db_xref="gi:22317004"
/translation="MGVPPNFWLPRLGLALPRLPAHYOFLDGVRLVSLTE
RGPVHSDSCPELTIRLRIPDPCPAPOIDRFQYDDEANARENAVGKALGFRT
GTMALCTVYKRGIAAGDALIEIRLRPDSIETVEQKAVFYQRTK"
CDS
BASE COUNT 259 a 247 c 233 g 139 t
ORIGIN

Query Match 99.5%; Score 703.8; DB 6; Length 878;
Best Local Similarity 99.7%; Pred. No. 2.1e-116;
Matches 705; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGACCCGGTGTCTGCTGCTTCCAGAGATGGGCGTGCACCCCACTTCCTCG 60
DB 59 TGACCCGGTGTCTGCTGCTTCCAGAGATGGGCGTGCACCCCACTTCCTCG 118
QY 61 GTGCTCCGGGCGGCTGGCGGAGTGGCGCTGCGCGGCTCCCGCCCACTACAGTTG 120
DB 119 GTGCTCCGGGCGGCTGGCGGAGTGGCGCTGCGCGGCTCCCGCCCACTACAGTTG 178
QY 121 CTGTTGACCTGGGCGTGGCGACCTGTGCTCCAGAGAGCGGCGGCGGCTCAGAC 180
DB 179 CTGTTGACCTGGGCGTGGCGACCTGTGCTCCAGAGAGCGGCGGCGGCTCAGAC 238
QY 181 GACAGCTCCCGGCGGCTGACCTGACCGGCTGCGGCTCCCGGCTTCGCGCGGCG 240
DB 239 GACAGCTCCCGGCGGCTGACCTGACCGGCTGCGGCTCCCGGCTTCGCGCGGCG 298
QY 241 CCGGACAGATGACCGGCTGTGCGAGATCGAGAGAGCGGCGGCGGAGAGGCT 300
DB 299 CCGGACAGATGACCGGCTGTGCGAGATCGAGAGAGCGGCGGCGGAGAGGCT 358
QY 301 GTGGAGTGCACCTGTGCTGCGGCTTTGGCGGCTGCGGCTGCGGCTGTTACCTG 360
DB 359 GTGGAGTGCACCTGTGCTGCGGCTTTGGCGGCTGCGGCTGCGGCTGTTACCTG 418
QY 361 GTGAAGAGCGGCGGCTTGGCTGCGAGAGATGCGATTCGTAATCCGAGACTACGACC 420
DB 419 GTGAAGAGCGGCGGCTTGGCTGCGAGAGATGCGATTCGTAATCCGAGACTACGACC 478
QY 421 GGCCTCATCAGAGACTATGAGAGAGAGAGAGAGTTCGATTCAGAGAGAGAGAA 480
DB 479 GGCCTCATCAGAGACTATGAGAGAGAGAGAGAGTTCGATTCAGAGAGAGAGAA 538
QY 481 TAAGGGGCTTAGTACCTTCTACAGGCGCTCACTCCCTTCCCATGTGTCATGGG 540
DB 539 TAAGGGGCTTAGTACCTTCTACAGGCGCTCACTCCCTTCCCATGTGTCATGGG 598
QY 541 GCCAGAGTGAAGGAGAGTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 600
DB 599 GCCAGAGTGAAGGAGAGTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 658
QY 601 CACTGAAGTACCCACCCCTGAGAGAGTCTGTATTAAGGAGAGGCTTGTACTGCTTT 660
DB 659 CACTGAAGTACCCACCCCTGAGAGAGTCTGTATTAAGGAGAGGCTTGTACTGCTTT 718
QY 661 GTGAATTAATGAGTTTACGACCAAAAAAAAAAAAAAAAAAAAAA 707
DB 719 GTGAATTAATGAGTTTACGACCAAAAAAAAAAAAAAAAAAAAAA 765

```

RESULT 3
LOCUS      AX099936      839 bp      DNA      linear      PAT 02-APR-2001
DEFINITION Sequence 18 from Patent WO0120004.
ACCESSION  AX099936
VERSION     AX099936.1  GI:13538946
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 839)
AUTHORS     Yue, H., Tang, Y. T., Bandman, O., Hillman, J. L., Baughn, M. R.,
             Azimzal, Y. and Lu, D. A.
TITLE       Protein phosphatase and kinase proteins
JOURNAL     Patent: WO 0120004-A 18 22-MAR-2001;
            Incyte Genomics, Inc. (US)
FEATURES
             location/Qualifiers
             source          1..839
                             /organism="Homo sapiens"
                             /db_xref="taxon:9606"
                             /note="Incyte ID NO: 1606974CH1"
BASE COUNT  161 a      258 c      255 g      165 t
ORIGIN
Query Match      99.4%; Score 703; DB 6; Length 839;
Best Local Similarity 100.0%; Pred. No. 2.9e-116;
Matches 703; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1  TGAACCCGCTCTCTGTCGCTTTCCTCCAGCATGCGCGTGAAGCCCACTTCTCTGG 60
DB      74  TGAACCCGCTCTCTGTCGCTTTCCTCCAGCATGCGCGTGAAGCCCACTTCTCTGG 133
OY      61  GTGCTTCGCGGGCGGGGGGAGAGTGGGCGTCCCGGCTCCCGCCACTACAGTTC 120
DB      134  GTGCTTCGCGGGCGGGGGGAGAGTGGGCGTCCCGGCTCCCGCCACTACAGTTC 193
OY      121  CTGTTGAGACTGGGCGTGCAGCAGTGTCTGCTGAGAGCGGCGGCGCCCTACAGC 180
DB      194  CTGTTGAGACTGGGCGTGCAGCAGTGTCTGCTGAGAGCGGCGGCGCCCTACAGC 253
OY      181  GACAGTGGCGGCGCTACCTGACCGCTGCGCATCCCGCATCTTCTGCGCGGCC 240
DB      254  GACAGTGGCGGCGCTACCTGACCGCTGCGCATCCCGCATCTTCTGCGCGGCC 313
OY      241  CCCGACAGATGAGCGCTTGTGAGATGTGAGAGAGGCGCAACGAGGAGAGGCT 300
DB      314  CCCGACAGATGAGCGCTTGTGAGATGTGAGAGAGGCGCAACGAGGAGAGGCT 373
OY      301  GTGGAGATGACATGTCTGTGGCTTGGCGGCTGAGCATGCTGCTGTACCTG 360
DB      374  GTGGAGATGACATGTCTGTGGCTTGGCGGCTGAGCATGCTGCTGTACCTG 433
OY      361  GTGAAGAGCGGGGCTTGGCTGAGAGATGCTTGTGAATCCGAGACTACAGCC 420
DB      434  GTGAAGAGCGGGGCTTGGCTGAGAGATGCTTGTGAATCCGAGACTACAGCC 493
OY      421  GGCTCATGAGAGCTATGAGAGGAGAAAGCTTCCAGTTCACAGCAAGCAAGAAA 480
DB      494  GGCTCATGAGAGCTATGAGAGGAGAAAGCTTCCAGTTCACAGCAAGCAAGAAA 553
OY      481  TAAAGGGCTTATGATCCCTTACAGGCGCTACCTCCCTCCATGTTGATGGG 540
DB      554  TAAAGGGCTTATGATCCCTTACAGGCGCTACCTCCCTCCATGTTGATGGG 613
OY      541  GCCAGAGATGAAGGAGAGTGAATTAAGTATTAACCTCTAGCTCCATTTGGCTGAAGA 600
DB      614  GCCAGAGATGAAGGAGAGTGAATTAAGTATTAACCTCTAGCTCCATTTGGCTGAAGA 673
OY      601  CACTGAAGTAGCCACCCCTGAGAGGAGGCTGATGAAGGAGGAGGCTTGTACTCTTT 660
DB      674  CACTGAAGTAGCCACCCCTGAGAGGAGGCTGATGAAGGAGGAGGCTTGTACTCTTT 733

```

```

OY      661  GTGATTAATGAGTTTACGACCAACCAAAAAAAAAAAAAAAAAA 703
DB      734  GTGATTAATGAGTTTACGACCAACCAAAAAAAAAAAAAAAAAA 776

RESULT 4
LOCUS      BC001140      711 bp      mRNA      linear      PRI 12-JUL-2001
DEFINITION Homo sapiens, hypothetical protein FLJ20442, clone MGC:1042
ACCESSION  BC001140
VERSION     BC001140.1  GI:12654608
KEYWORDS
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 711)
AUTHORS     Strausberg, R.
TITLE       Direct Submission
JOURNAL     Submitted (11-DEC-2000) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA

```

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgaps-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcsc.bc.ca

Steven Jones, Jennifer Asano, Ian Boscut, Yaron Butterfield, Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandon, Anna-Lisa Prabhu, Parvaneh Saedi, Jacqueline Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRL Plate: 8 Row: 0 Column: 23
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7020544.

FEATURES
 source
 location/Qualifiers

```

1..711
/organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="taxon:9606"
/clone="MGC:1042 IMAGE:3535406"
/tissue_type="Lung, small cell carcinoma"
/clone_lib="NHL MGC-7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
21..473
/codon_start=1
/product="hypothetical protein FLJ20442"
/protein_id="AAH01140.1"
/db_xref="GI:12654608"
/translation="MGVPPNPSVLPRLAGLALPRLPAHYQFLLDGVRLVSLTE
RGPSPDSGPGITLRLRIIPDCPPAPDIDRFVQIVDEANRGAVGVHCAIGGRT
GTMLACYLVKKGGLAAGDAIAEIRLRGSIETRYOEKAVPQFVYRTK"

```

BASE COUNT 159 a 217 c 208 g 127 t

Query Match 96.2%; Score 679.8; DB 9; Length 711;
 Best Local Similarity 99.7%; Pred. No. 4.1e-112;
 Matches 681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY	25	CCAGGATGGGGGTGCAGGCCCGCCCAACTCTCTGGGTGTTCCGGGGCGGGTGGGGGA	84
Dd	15	CCAGCGATGGGGGTGCAGGCCCGCCCAACTCTCTGGGTGTTCCGGGGCGGGTGGGGGA	74
OY	85	CTGGCGCTGCCCGGCTCCCCGCCACATACAGTTCTCTTGAGACCCTGGGCTGGCGCAC	144
Dd	75	CTGGCGCTGCCCGGCTCCCCGCCACATACAGTTCTCTTGAGACCCTGGGCTGGCGCAC	134
OY	145	CTGGTCTCCTTCACGAGAGGGGGGGCCCTTCACACGAGACAGTGCCTCGGCTCACCTTG	204
Dd	135	CTGGTCTCCTTCACGAGAGGGGGGGCCCTTCACAGGAGACGCTGCCCGCCCTCACCCG	194
OY	205	CACCCCTCTGCGATCCCGCATCTTGCCCGCGGGCCCCGACAGATGACCGCTTCG	264
Dd	195	CACCCCTCTGCGATCCCGCATCTTGCCCGCGGGCCCCGACAGATGACCGCTTCG	254
OY	265	CAGATCTGTGAGAGAGGCCAACGACAGGGGAGAAGCTGTGGAGTAGCATGTCTTGAGG	324
Dd	255	CAGATCTGTGAGAGAGGCCAACGACAGGGGAGAAGCTGTGGAGTAGCATGTCTTGAGG	314
OY	325	TTTGGCCCGACATCGGACACATCTGTGGCTGTTCCTGTGTGAAGAGCGGGGCTTGGCTGA	384
Dd	315	TTTTGGCCCGACATCGGACACATCTGTGGCTGTTCCTGTGTGAAGAGCGGGGCTTGGCTGA	374
OY	385	GGAGATGCGCATCTGTGAATCCGAGAGCTACGACCCGGCTCCATGAGACCTATGAGCAG	444
Dd	375	GGAGATGCGCATCTGTGAATCCGAGAGCTACGACCCGGCTCCATGAGACCTATGAGCAG	434
OY	445	GAGAAAGCAGCTCTCCAGTTCTACGAGGAAAGAAATAAAGGGCCCTTAATACCTCTTAC	504
Dd	435	GAGAAAGCAGCTCTCCAGTTCTACGAGGAAAGAAATAAAGGGCCCTTAATACCTCTTAC	494
OY	505	CAGGCCCTACCTCCCTTCCCATTGTGTGATGGGGCCAGAGATGAAGGGAAGTGGACT	564
Dd	495	CAGGCCCTACCTCCCTTCCCATTGTGTGATGGGGCCAGAGATGAAGGGAAGTGGACT	554
OY	565	AAAGATTAAACCCCTCTAGCTCCCATTTGGCTTAAGACACTGAAGTGAAGCCACCTGGAG	624
Dd	555	AAAGATTAAACCCCTCTAGCTCCCATTTGGCTTAAGACACTGAAGTGAAGCCACCTGGAG	614
OY	625	GCAGTCTCTGATTGAAGGGGAGGCTTGATGCTTGTGTAATGAATGATTTTACGAAC	684
Dd	615	GCAGTCTCTGATTGAAGGGGAGGCTTGATGCTTGTGTAATGAATGATTTTACGAAC	674
OY	685	CAAAAAAAAAAAAAAAAAAAAAA 707	
Dd	675	CAGGAAAAAAAAAAAAAAAAAAAA 697	

RESULT 5
AX398840

LOCUS AX398840 746 bp DNA linear PAT 27-MAY-2002

DEFINITION Sequence 11 from Patent WO0220747.

ACCESSION AX398840

VERSION AX398840.1 GI:21261352

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

AUTHORS Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE Regulation of human tyrosine phosphatase-like enzyme

JOURNAL Patent: WO 0220747-A 11 14-MAR-2002;

FEATURES

source Location/Qualifiers

1..746 /organism="Homo sapiens"

BASE COUNT 178 a 221 c 218 g 129 t

Query Match 95.3%; Score 673.6; DB 6; Length 746;

Best Local Similarity 99.4%; Pred. No. 5.3e-111;									
Matches 676; Conservative 0; Mismatches 4; Indels 0; Gaps 0;									
OY	28	GGGATGGGCGTCCAGCCGCCCAACTTCTCTGGGTCTCCGGGCCGGCTGGCGGACTG	87						
DB	32	GGGAATGGGGGTGCACACCCCCCAACTTCTCTGGGTCTCCGGGCCGGCTGGCGGACTG	91						
OY	88	GGGCTGGCGGGCTCCCGGCCCACTACAGATCTCTGGTGGACTGGGGGTGGGGACCTG	147						
DB	92	GGGCTGGCGGGCTCCCGGCCCACTACAGATCTCTGGTGGACTGGGGGTGGGGACCTG	151						
OY	148	GTGTCCCTGACGAGAGGGGGGCCCTCAGACAGAGCTGGCCCGGCTCACCTCGAC	207						
DB	152	GTATCCCTGAGAGGAGGGGGGCCCTCAGACAGAGCTGGCCCGGCTCACCTCGAC	211						
OY	208	CGCTTGGCGATCCCGGACTTCTGCCCCGGCCGCCAGACAGATGACCGCTTCGTGAC	267						
DB	212	CGCTTGGCGATCCCGGACTTCTGCCCCGGCCGCCAGACAGATGACCGCTTCGTGAC	271						
OY	268	ATCGTGGAGAGGAGGCAAGGACGGGGAGAGCGCTGGGAGTGCATGTGCTTGGGCTT	327						
DB	272	ATCGTGGAGAGGAGGCAAGGACGGGGAGAGCGCTGGGAGTGCATGTGCTTGGGCTT	331						
OY	328	GGCCGCGACTGGCAGCATGCTGGCTGTACCTGTGTGAAGAGAGCGGGCTTGGCTCAGAG	387						
DB	332	GGCGCGACTGGGAGCATGCTGGCTGTACCTGTGTGAAGAGAGCGGGCTTGGCTCAGAG	391						
OY	388	GATGCCATTGCTGAATCCGACGACTACGACCCGGCTCCATGAGACCTATGAGCAGAG	447						
DB	392	GATGCCATTGCTGAATCCGACGACTACGACCCGGCTCCATGAGACCTATGAGCAGAG	451						
OY	448	AAAGAGCTTCCCAATTTACAGAGGAAGAAATAAGGGGCTTGTAGTACCTTTACACAG	507						
DB	452	AAAGAGCTTCCCAATTTACAGAGGAAGAAATAAGGGGCTTGTAGTACCTTTACACAG	511						
OY	508	GCCCTCACTCCCTTCCCATGTTGTGCATGGGGGCGCAAGATGAGAGGGAAGTGACTAA	567						
DB	512	GCCCTCACTCCCTTCCCATGTTGTGCATGGGGGCGCAAGATGAGAGGGAAGTGACTAA	571						
OY	568	GTAATTAACCCCTCTAGCTGCCATTGGCTGAAGACACTGAAGGCCACCCCTGCAGCA	627						
DB	572	GTAATTAACCCCTCTAGCTGCCATTGGCTGAAGACACTGAAGGCCACCCCTGCAGCA	631						
OY	628	GGTCTGTATGAAGGGGAGGCTGTAGTACGCTTTGTGTAATAATGATTTTACAGACCA	687						
DB	632	GGTCTGTATGAAGGGGAGGCTGTAGTACGCTTTGTGTAATAATGATTTTACAGAACCA	691						
OY	688	AAAAAAAAAAAAAAAAAAAAA 707							
DB	692	GGAAAAAAAAAAAAAAAAAAAAA 711							
RESULT 6									
LOCUS	AX398843	746 bp	DNA linear PAT 27-MAY-2002						
DEFINITION	Sequence 14 from Patent W00220747.								
ACCESSION	AX398843								
VERSION	AX398843.1 GI:21261353								
KEYWORDS	human.								
SOURCE	human.								
ORGANISM	Homo sapiens								
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.								
AUTHORS	Kossida, S.								
TITLE	1 Regulation of human tyrosine phosphatase-1-like enzyme								
JOURNAL	Patent: WO 0220747-A 14 14-MAR-2002;								
FEATURES	Bayer Aktiengesellschaft (DE)								
source	Location/Qualifiers								
	1..746								
	/organism="Homo sapiens"								
	/db_xref="taxon:9606"								
BASE COUNT	178 a 221 c 218 g 129 t								

AUTHORS Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T., Suzuki Y., Ohashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 746)
 AUTHORS Sugano S., Suzuki Y., Ota T., Ohashi M., Nishi T., Isogai T., Shibahara T., Tanaka T., and Nakamura Y.
 TITLE Direct Submission
 JOURNAL Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-da, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'-3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES
 source 1..746
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="KAT04828"
 /cell_line="KATO III"
 /cell_type="signet-ring cell carcinoma"
 /clone_1db="KAT"
 /note="cloning vector pME18SFL3"
 35..487
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BA091172.1"
 /db_xref="GI:7020545"
 /translation="MGVPPNFSNVLPRLGALPLPAHYOFLIDGVRHYVSLTE RSPHSDSCPGTLRLRIPDRCPAPDQIDRFQYIDENAKREAYGVKALFGRGT GYMALCTIVERGIAAGDALAEIRLRPGLPTEYQEKAVFOYRRTK"

CDS
 178 a 221 c 218 g 129 t

BASE COUNT
 178 a 221 c 218 g 129 t

ORIGIN

Query Match 95.3%; Score 673.6; DB 9; Length 746;
 Best Local Similarity 99.4%; Pred. No. 5.3e-111;
 Matches 676; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

28 GCGATGGCGCTGACAGCCCCCAACTCTCTGCGTCTCCGGCGCGGCTGCGGAGCTG 87
 |||||
 Db 32 GCGATGGCGCTGACAGCCCCCAACTCTCTGCGTCTCCGGCGCGGCTGCGGAGCTG 91
 |||||
 QY 88 GCGCTGCGGCGCTGCGGCGGCTGCGGAGCTGCGGCGGCTGCGGAGCTG 147
 |||||
 Db 92 GCGCTGCGGCGCTGCGGCGGCTGCGGAGCTGCGGCGGCTGCGGAGCTG 151
 |||||
 QY 148 GTGTCCCTGAGAGGCGGCGGCGGCTGCGGAGCTGCGGCGGCTGCGGAGCTG 207
 |||||
 Db 152 GTGTCCCTGAGAGGCGGCGGCGGCTGCGGAGCTGCGGCGGCTGCGGAGCTG 211
 |||||
 QY 208 GCGCTGCGGCGCTGCGGAGCTGCGGCGGCGGCTGCGGAGCTGCGGAGCTG 267
 |||||
 Db 212 GCGCTGCGGCGCTGCGGAGCTGCGGCGGCGGCTGCGGAGCTGCGGAGCTG 271
 |||||
 QY 268 ATCTGTGAGAGGCGGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTG 327
 |||||
 Db 272 ATCTGTGAGAGGCGGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTG 331
 |||||
 QY 328 GCGCGGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTG 387
 |||||
 Db 332 GCGCGGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTG 391
 |||||
 QY 388 GATGCGCATGCTGAATCCGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 447
 |||||
 Db 392 GATGCGCATGCTGAATCCGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 451
 |||||
 QY 448 AAAGCAGCTTTCAGATTCTACAGAGAAATAAAGGCGCTTAGTACCTTACACAG 507
 |||||

Db 452 AAAGCAGCTTTCAGATTCTACAGAGAAATAAAGGCGCTTAGTACCTTACACAG 511
 |||||
 QY 508 GCGCGGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTG 567
 |||||
 Db 512 GCGCGGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTG 571
 |||||
 QY 568 GATTTAAACCTCTAGCTCCATGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 627
 |||||
 Db 572 GATTTAAACCTCTAGCTCCATGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 631
 |||||
 QY 628 GCTCTGATGAGAGGAGGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 687
 |||||
 Db 632 GCTCTGATGAGAGGAGGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 691
 |||||
 QY 688 AAAAAAAAAAAAAAAAAA 707
 |||||
 Db 692 GAAAAAAAAAAAAAAAAA 711
 |||||

RESULT 9
 AX398833 709 bp DNA linear PAT 27-MAY-2002
 LOCUS AX398833
 DEFINITION Sequence 4 from patent WO0220747.
 ACCESSION AX398833
 VERSION AX398833.1 GI:21261345
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Kossida S.
 TITLE Regulation of human tyrosine phosphatase-like enzyme
 JOURNAL Patent: WO 0220747-A 4 14-MAR-2002;
 Bayer Aktiengesellschaft (DE)
 FEATURES
 source 1..709
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT
 137 a 233 c 206 g 133 t

ORIGIN

Query Match 95.0%; Score 671.6; DB 6; Length 709;
 Best Local Similarity 98.6%; Pred. No. 1.2e-110;
 Matches 688; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 1 TGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
 |||||
 Db 13 TGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 72
 |||||
 QY 61 GTGCTTCGGGCGCGGCTGCGGAGCTGCGGCGGCTGCGGCGGCTGCGGAGCTG 120
 |||||
 Db 73 GTGCTTCGGGCGCGGCTGCGGAGCTGCGGCGGCTGCGGCGGCTGCGGAGCTG 132
 |||||
 QY 121 CTGTGTGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTG 180
 |||||
 Db 133 CTGTGTGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTG 192
 |||||
 QY 181 GACAGCTGCCCGGCTGACCTGACCGCTGACCGCTGACCGCTGACCGCTGACCG 240
 |||||
 Db 193 GACAGCTGCCCGGCTGACCTGACCGCTGACCGCTGACCGCTGACCGCTGACCG 252
 |||||
 QY 241 CCCGACGAGATGACAGCTGCTGCTGAGATGCTGAGAGGAGGAGGAGGAGGAG 300
 |||||
 Db 253 CCCGACGAGATGACAGCTGCTGCTGAGATGCTGAGAGGAGGAGGAGGAGGAG 312
 |||||
 QY 301 GTGGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 360
 |||||
 Db 313 GTGGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 372
 |||||
 QY 361 GTGAAGAGCGGCGCTGCTGAGAGAGTGCATGTGTAATCCAGACTAGACCC 420
 |||||

Db 373 GTGAAGGAGCGGGGCTGGCTGCAGAGAGATCCATTGCTGTAATCCGACGATACGAGCC 432
QY 421 GGCTCATCGAGACCTATGACGAGAGAAAGACAGTCTTCAGTTCTTACAGGAAAGAA 480
Db 433 GGCTCATCGAGACCTATGACGAGAGAAAGACAGTCTTCAGTTCTTACAGGAAAGAA 492
QY 481 TAAAGGGCCCTTAAATGACCTTCTTACAGGAGCCCTCACTCCCTCCCTCCATGTTGATGG 540
Db 493 TAAAGGGCCCTTAAATGACCTTCTTACAGGAGCCCTCACTCCCTCCCTCCATGTTGATGG 552
QY 541 GCCAGATGAGAGGAGTGTGACATTAAGTATTAACCTCTTACGTTCCATTGGCTGAAGA 600
Db 553 GCCAGATGAGAGGAGTGTGACATTAAGTATTAACCTCTTACGTTCCATTGGCTGAAGA 612
QY 601 CACTGAGTACCCACCCCTGCAGGAGGCTCTGATTGAAGGAGGCTGTACTGCTTT 660
Db 613 CACTGAGTACCCACCCCTGCAGGAGGCTCTGATTGAAGGAGGCTGTACTGCTTT 671
QY 661 GTTGAATTAATGATTTTACGACCAAAAAAAAAAAAA 698
Db 672 GTCAATTAATGATTTTACGACCAAAAAAAAAAAAA 709

RESULT 10

AX086042 694 bp DNA linear PAT 09-MAR-2001
LOCUS Sequence 39 from Patent WO0112819.
DEFINITION AX086042
ACCESSION AX086042.1 GI:13275852
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 694)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Protein phosphatases and diagnosis and treatment of
JOURNAL Patient: WO 0112819-A 39 22-FEB-2001;
Sugen, Inc. (US)

FEATURES

source 1..694
Location/Qualifiers

BASE COUNT 128 a 221 c 216 g 129 t
ORIGIN
Query Match 93.0%; Score 657.4; DB 6; Length 694;
Best Local Similarity 99.8%; Pred. No. 4.2e-108;
Matches 658; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 28 GCATGAGGCGTGCAGCCGCCAATTCTCGTGGTTCGCGCCGCGCGGAGACTG 87
Db 32 GCATGAGGCGTGCAGCCGCCAATTCTCGTGGTTCGCGCCGCGCGGAGACTG 91
QY 88 GCCTGCGCGGCGTCCCGCCCACTACAGTCTTGTGACCGGCGTGGGACACTG 147
Db 92 GCCTGCGCGGCGTCCCGCCCACTACAGTCTTGTGACCGGCGTGGGACACTG 151
QY 148 GTTCTCCCTGACGAGCGCGGCGCCCTCAGACGAGACGTGCCCGGCTCACCTGAC 207
Db 152 GTTCTCCCTGACGAGCGCGGCGCCCTCAGACGAGACGTGCCCGGCTCACCTGAC 211
QY 208 CGCTGCGGCGTCCCGCGCTTCGCGCGCGCGCCGACAGATCCAGCCGCTTGTGAC 267
Db 212 CGCTGCGGCGTCCCGCGCTTCGCGCGCGCGCCGACAGATCCAGCCGCTTGTGAC 271
QY 268 ATGCTGAGAGGCGCAACGACGCGGAGAGGCTGTGGAGTGTGACTGTGCTGGGCTTT 327
Db 272 ATGCTGAGAGGCGCAACGACGCGGAGAGGCTGTGGAGTGTGACTGTGCTGGGCTTT 331
QY 328 GAGCGGAGTGTGACGATGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 387

Db 332 GCGCGACATGCGACATGCTGCTGTTACCTGTGGAAGAGCGGGGCTTGGCTGCAGGA 391
QY 388 GATGCCATTTGCTGAATTCGACGACACTACGACCCGCTTCATGAGACCTATGAGCAGAG 447
Db 392 GATGCCATTTGCTGAATTCGACGACACTACGACCCGCTTCATGAGACCTATGAGCAGAG 451
QY 448 AAAGCAGTCTTCAGTCTTACACGAGAAATTAAGGGGCTTGTGATCCCTTACACAG 507
Db 452 AAAGCAGTCTTCAGTCTTACACGAGAAATTAAGGGGCTTGTGATCCCTTACACAG 511
QY 508 GCCCTCACTCCCTCCCATGTTGTGATGGGCGCAGAGATGAAGGGAAGTGAAGTAA 567
Db 512 GCCCTCACTCCCTCCCATGTTGTGATGGGCGCAGAGATGAAGGGAAGTGAAGTAA 571
QY 568 GTATTAACCTCTTACGTTCCCATTTGCTGTAAGACACTGAATACCCACCCCTCAGGCA 627
Db 572 GTATTAACCTCTTACGTTCCCATTTGCTGTAAGACACTGAATACCCACCCCTCAGGCA 631
QY 628 GGTCTGATGAAGGGGAGGCTGTGCTGTTGTAATTAAGATTTACGACCA 686
Db 632 GGTCTGATGAAGGGGAGGCTGTGCTGTTGTAATTAAGATTTACGACCA 690

RESULT 11

AX398836 602 bp DNA linear PAT 27-MAY-2002
LOCUS Sequence 7 from Patent WO0220747.
DEFINITION AX398836
ACCESSION AX398836.1 GI:21261348
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Kossida, S.
TITLE Regulation of human tyrosine phosphatase-like enzyme
JOURNAL Patient: WO 0220747-A 7 14-MAY-2002;
Bayer Aktiengesellschaft (DE)

FEATURES

source 1..602
Location/Qualifiers

BASE COUNT 109 a 202 c 179 g 110 t 2 others
ORIGIN
Query Match 82.6%; Score 583.8; DB 6; Length 602;
Best Local Similarity 99.3%; Pred. No. 6.5e-95;
Matches 596; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 42 GCGCCCACTTCTCTGCTGCTTCCGCGCGGCGTGGCGAGCTGGCGGCTCGCGGCT 101
Db 3 GCGCCCACTTCTCTGCTGCTTCCGCGCGGCGTGGCGAGCTGGCGGCTCGCGGCT 62
QY 102 CCGCGCCACTTACAGTCTGTTGACGTGGGCGTGGCGGCTGCTGCTGCTGCTGCTG 161
Db 63 CCGCGCCACTTACAGTCTGTTGACGTGGGCGTGGCGGCTGCTGCTGCTGCTGCTG 122
QY 162 GCGCGGCGGCGTCCACAGCAGACTGCCCGGCTTCACCTGACAGCGCTGGCATGCC 221
Db 123 GCGCGGCGGCGTCCACAGCAGACTGCCCGGCTTCACCTGACAGCGCTGGCATGCC 182
QY 222 CGACTTCTGCGCGCGCGCGCCGACAGATGACCGCTTGTGAGTGTGAGCAGAGC 281
Db 183 CGACTTCTGCGCGCGCGCGCCGACAGATGACCGCTTGTGAGTGTGAGCAGAGC 242
QY 282 CAACGACGAGGAGAGGCTGTGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 341
Db 243 CAACGACGAGGAGAGGCTGTGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 302
QY 342 CATGCTGCTGCTTACCTGTGTAAG-GAAGCGGCGCTTGGCTGAGAGATGATGCTG 400

Db	303	CATCTGTGGCTGTACTGTGTGAANGAGGGGNCCTTGGTGCAGAGAGATGCCATTCTG	362
QY	401	AAATCCGACGACTCTACACACCCGGCTCTCATCGAGACTTGTAGCAGAGAAAGCACTTCC	460
Db	363	AAATCCGAGCATACGACACCCGGCTCATCGAGACTGTGTAGCAGAGAAAGCACTTCTCC	422
QY	461	AGTTCTACGACGAGCAACAATTAAGGGGCGCTTAGTACCTCTTCTACAGGCCCTCACTCCC	520
Db	423	AGTTCTACGACGAGCAACAATTAAGGGGCGCTTAGTACCTCTTCTACAGGCCCTCACTCCC	483
QY	521	TTCCGCCATGTTGTGTGATGAGGGGCCAGAGATGAGGGAAGTGAGCTAAAGTATTAACCCCTC	580
Db	483	TTCCGCCATGTTGTGTGATGAGGGGCCAGAGATTAAGGGAAGTGAGCTAAAGTATTAACCCCTC	542
QY	581	TACGTCCCATTTGGCGTGAAGACATCGAAGTATGCCCCACCTCGCAGGGCGAGTCCGATTGAA	640
Db	543	TAGTCCCATTTGGCGTGAAGACATCGAAGTATGCCCCACCTCTCGAGGCGAGTCTCGATTGAA	602

RESULT	12
HUMZE12H05	
LOCUS	
DEFINITION	559 bp mRNA linear.
DESCRIPTION	Homo sapiens full length insert cDNA clone ZEL12H05.
ACCESSION	M606657
PRI	29-AUG-1998

SOURCE ORGANISM

REFERENCE AUTHORS

TITLE	Full Clone Sequencing of the Longest Available Member from Each Unigene Cluster
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

JOURNAL	unpublished	
REFERENCE	2 (bases 1 to 559)	
AUTHORS	Waterston, R.	
TITLE	Direct Submission	
JOURNAL	Submitted (24-ADG-1998)	Department of Genetics, Washington
	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA	
COMMENT	SUBMITTED BY:	

Genome Sequencing Center
Department of Genetics
Washington University
St. Louis MO 63108, USA
<http://genome.wustl.edu/gsc>
<mailto:est@watson.wustl.edu>

NOTICE: This sequence represents the full insert of this cDNA. No attempt has been made to verify whether this corresponds to the full-length of the original mRNA from which it was derived. We have tried to obtain double-stranded, or double chemistry sequence across the entire clone, but potentially, there are areas in the sequence where this level of coverage was not achieved. Nevertheless, we are confident of the accuracy of this sequence as all regions of low quality, as defined by PhRAP (P. Green, in preparation), were visually inspected and edited accordingly. The consensus quality values for this sequence have been submitted separately.

SIMILARITY INFORMATION:

similar to Clona intestinalis protein Z83760 (PID:g1764017) COS41.7

The location of this clone is unknown.

FEATURES
source

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"
```

BASE COUNT	ORIGIN
121 a	169 c 162 g 107 t
misc_feature	/clone="IMAGE:358809" /clone_lib="Soares_fetal_heart_NbHH19w"
misc_feature	54. .339 /note="Similar to Clona intestinalis protein 283760 (PID:g1764017)"
misc_feature	61. .330 /note="Similar to Homo sapiens protein AF064105 (PID:g3136334)"
misc_feature	85. .333 /note="Similar to Pyrococcus horikoshii protein AB009524 (PID:g3132001)"
misc_feature	91. .287 /note="Similar to Caenorhabditis elegans protein U28739 (PID:g2733378)"
misc_feature	91. .297 /note="Similar to Caenorhabditis elegans protein U28739 (PID:g2733379)"

Query Match	78.6%;	Score 555.8;	DB 9;	Length 559;
Best Local Similarity	-99.6%;	Pred. No. 6.7e-90;		
Matches 557; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY 133 GGGGTGGCGCACCTGTGTCCTTGACGAGAGCGGGGCCCCCTACAGCAGCAGTGGCCCC 192

Db 1 GGGGTGGCGCACCTGTGTCCTTGACGAGAGCGGGGCCCCCTACAGCAGCAGTGGCCCC 60

QY	GGCTCACCCTGCACCGCCCTGCGCATCCCCGACTTCTGCGCGCGCGCCCCGACAGATC	252
193	GGCTCACCCTGCACCGCCCTGCGCATCCCCGACTTCTGCGCGCGCGCCCCGACAGATC	
db	GGCTCACCCTGCACCGCCCTGCGCATCCCCGACTTCTGCGCGCGCGCCCCGACAGATC	
61	GGCTCACCCTGCACCGCCCTGCGCATCCCCGACTTCTGCGCGCGCGCCCCGACAGATC	120

253 GACCGCTTCGTGCAGATCGTGGACGAGGCCAACGCACGGGGAGAGGCTGTGGAGATCCAC 312

313 TGTGCTCTGGGCTTTGGCCGCACTGGCACCACCTGGGCGCTGTAACTGGTGAAGACCG 372

373 GCGTTGGCTGCAGGAGATGCCATTGCTCGAATTCGACGACTACGACCCGGCTCCATCGAG 432
 DB 181 TGTGCTCTGCGGCTTTGGCCCGGACATGGCACCATCTGGCCCTGTTACTCTGGTGAAGGAGCGG 240
 OY

[illegible]

Db 301 ACCATGAGCAGGAGAAAGCAGTCTTCCAGTTCTACCGACGACACAATAAGGGCCCTTA 360

Db 361 GTACCCCTTACCAAGGCCCTCACTCCCCCTGCCCATGTTGCGATGGGGGCCAGAGATGAA 420

QY	GGGAGTGGACTAAAGTATTAAACCCCTAGCTCCCATTTGGCTGAAGACACTGAAGTATGC	612
553	GGGAGTGGACTAAAGTATTAAACCCCTAGCTCCCATTTGGCTGAAGACACTGAAGTATGC	
421	GGGAGTGGACTAAAGTATTAAACCCCTAGCTCCCATTTGGCTGAAGACACTGAAGTATGC	480

Qy 613 CCACCCCGACGAGGAGCTCTGATTTGAAGGGAGGCTTGTAAGCTTTGTTGAATAAATG 672
|||||
A81 CCCCCCAGGAGGAGGCTCTGATTTGAAGGGAGGCTTGTAAGCTTTGTTGAATAAATG 672
|||||

QY 673 AGTTTACGACCAAAAA 691
||||| 111111

RESUME 13

LOCUS	AX398838	553 bp	DNA	linear	PAT 27-MAY-2007
DEFINITION	Sequence 9 from Patent WO0220747.				

```

KEYWORDS    AX398838.1  GI:21261350
VERSION
SOURCE      human.

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Kossida, S.
JOURNAL Regulation of human tyrosine phosphatase-like enzyme
Patent: WO 0220747-A 9 14-MAR-2002;
Bayer Aktiengesellschaft (DE)
Location/Qualifiers
source 1..353
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 96 a 186 c 168 g 103 t
ORIGIN

Query Match 71.9%; Score 508.6; DB 6; Length 553;
Best Local Similarity 98.6%; Pred. No. 1.9e-81;
Matches 545; Conservative 0; Mismatches 4; Indels 4; Gaps 3;

QY 29 CGATGGCGGTGACACCCCAACTTCTCTGGGTGCTTCCGGGCGGCTGGCGGACTGG 88
DB 1 CGATGGCGGTGACACCCCAACTTCTCTGGGTGCTTCCGGGCGGCTGGCGGACTGG 60
QY 89 CGTGGCGGCGGTGACACCCCAACTTCTCTGGGTGCTTCCGGGCGGCTGGCGGACTGG 148
DB 61 CGCTGGCGGCGGTGACACCCCAACTTCTCTGGGTGCTTCCGGGCGGCTGGCGGACTGG 120
QY 149 TGTCCCTGAC--GGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 206
DB 121 TGTCCCTGACGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 207 CCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 265
DB 181 CCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 266 AGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 324
DB 241 AGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 325 TTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 384
DB 301 TTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 360
QY 385 GGAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 444
DB 361 GGAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY 445 GAGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 504
DB 421 GAGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 505 CAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 564
DB 481 CAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
QY 565 AAAGTATTAAC 577
DB 541 AAAGTATTAAC 553

RESULT 14
AX398834 493 bp DNA 11linear PAT 27-MAY-2002
LOCUS
DEFINITION Sequence 5 from Patent WO0220747.
ACCESSION AX398834
VERSION AX398834.1 GI:21261346
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Kossida, S.

TITLE Regulation of human tyrosine phosphatase-like enzyme
JOURNAL Patent: WO 0220747-A 5 14-MAR-2002;
Bayer Aktiengesellschaft (DE)
Location/Qualifiers
source 1..493
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 78 a 176 c 149 g 90 t
ORIGIN

Query Match 68.8%; Score 486.2; DB 6; Length 493;
Best Local Similarity 99.4%; Pred. No. 2e-77;
Matches 488; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 42 GCCCGCACTTCTCTGGGTGCTTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 101
DB 3 GCCCGCACTTCTCTGGGTGCTTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 62
QY 102 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 161
DB 63 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 122
QY 162 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 221
DB 123 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 182
QY 222 GCACTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 281
DB 183 GCACTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 242
QY 282 CAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 341
DB 243 CAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 302
QY 342 CATGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 401
DB 303 CATGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 362
QY 402 AATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 461
DB 363 AATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 422
QY 462 GTTCTACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 521
DB 423 GTTCTACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 482
QY 522 TCCCGATGTTG 532
DB 483 TCCCGATGTTG 493

RESULT 15
AX398839/c 448 bp DNA 11linear PAT 27-MAY-2002
LOCUS
DEFINITION Sequence 10 from Patent WO0220747.
ACCESSION AX398839
VERSION AX398839.1 GI:21261351
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Kossida, S.
JOURNAL Regulation of human tyrosine phosphatase-like enzyme
Patent: WO 0220747-A 10 14-MAR-2002;
Bayer Aktiengesellschaft (DE)
Location/Qualifiers
source 1..448
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 93 a 132 c 120 g 103 t
ORIGIN

Query Match 63.1%; Score 446; DB 6; Length 448;
Best Local Similarity 100.0%; Pred. No. 3.1e-70;
Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	241	CCCGACCGAGATCGACCGCTTCGTGAGATCGTGAGAGAGGCCAACGACGAGGGAGAGGCT	300
Db	448	CCCGACCGAGATCGACCGCTTCGTGAGATCGTGAGAGAGGCCAACGACGAGGGAGAGGCT	389
QY	301	GTGGAGATGCACTGTGCTCTGGGGCTTTGGCCGACCTGGCACCATGTGGCTGTACCTG	360
Db	388	GTGGAGATGCACTGTGCTCTGGGGCTTTGGCCGACCTGGCACCATGTGGCTGTACCTG	329
QY	361	GTGAAGAGCGGGGCTTGGCTGCGAGAGATGCCATTGCTGAATCGAAGACTACGACCC	420
Db	328	GTGAAGAGCGGGGCTTGGCTGCGAGAGATGCCATTGCTGAATCGAAGACTACGACCC	269
QY	421	GGCTCCATCGAGACCTATGAGAGAGAAAGAGTCTCCAGTCTACACGCAACGAAA	480
Db	268	GGCTCCATCGAGACCTATGAGAGAGAAAGAGTCTCCAGTCTACACGCAACGAAA	209
QY	481	TAAGGGGCTTAGTACCTTCTACAGGCGCTCACTCCGCTTCCCATGTTGCGATGGG	540
Db	208	TAAGGGGCTTAGTACCTTCTACAGGCGCTCACTCCGCTTCCCATGTTGCGATGGG	149
QY	541	GCCAGAGATGAAGGAGTGACTTAAGTATTAAACCTCTAGCTCCATTGGCTGAAGA	600
Db	148	GCCAGAGATGAAGGAGTGACTTAAGTATTAAACCTCTAGCTCCATTGGCTGAAGA	89
QY	601	CACGTAAGTACCCCAACCCCTGCGAGGAGTCTGATTTGAAGGGAGGCTTGTACTGCTT	660
Db	88	CACGTAAGTACCCCAACCCCTGCGAGGAGTCTGATTTGAAGGGAGGCTTGTACTGCTT	29
QY	661	GTTGATATAATGAGTTTACGAAACA	686
Db	28	GTTGATATAATGAGTTTACGAAACA	3

Search completed: November 4, 2002, 08:49:28
Job time : 2789 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 2002, 06:37:01 : Search time 258 Seconds
(without alignments)
6171.174 Million cell updates/sec

Title: US-09-619-380-1

Perfect score: 707
Sequence: 1 tgaaccgctgtcctgtgcc.....aaaaaaaaaaaaaaaa 707

Scoring table: IDENTITY_NUC
Gap 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N_Geneseq_101002:*

```
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	707	100.0	707	22	AAF27958
2	703	99.4	839	22	AAF30482
3	673.6	95.3	746	24	ABL40814
4	673.6	95.3	746	24	ABL40815
5	671.6	95.0	709	22	ABL40807
6	657.4	93.0	694	22	AAF63583
7	583.8	82.6	602	24	ABL40810
8	508.6	71.9	553	24	ABL40812
9	486.2	68.8	493	24	ABL40808

10	471.8	66.7	559	24	AB056155	Human ovarian anti
11	446	63.1	448	24	ABL40813	Human tyrosine pho
12	429	60.7	453	24	ABL40809	Human tyrosine pho
13	396.4	56.1	715	22	AAF63563	Murine phosphatase
14	379.8	53.7	397	24	ABL40811	Human tyrosine pho
15	373	52.8	453	22	AAF27959	Murine dual specif
16	336.4	47.6	399	24	ABL40806	Human tyrosine pho
17	241.8	34.2	247	24	ABL67099	Thyroid cancer rel
18	60	8.5	60	24	ABN34834	Human spliced tran
19	55.2	7.8	1146	16	AA083733	S. avermitilis BCK
20	55.2	7.8	2728	16	AA083737	S. avermitilis bkd
21	49.6	7.0	1174	21	AA044085	Zea mays DNA fragm
22	49.6	7.0	1373	21	AA046590	Zea mays DNA fragm
23	49.2	7.0	1169	22	AA050864	Human cDNA encodin
24	49.2	7.0	1720	24	ABN85027	Human cDNA encodin
25	49.2	7.0	2936	22	AA080866	Human Tumour Suppr
26	49.2	7.0	3098	21	AA043459	Human cDNA encodin
27	49.2	7.0	3844	22	AA076202	CDNA encoding huma
28	49.2	7.0	3910	22	AA015815	Human drug metabol
29	49.2	7.0	3912	22	AA014598	Human cDNA encodin
30	48.8	6.9	35133	23	AB050991	Human cDNA sequenc
31	48.8	6.9	35134	23	AB050990	Thermus caldophilu
32	48.6	6.9	88421	24	AA040781	Thermus caldophilu
33	48	6.8	573	23	AB018649	8842int genomic DN
34	48	6.8	2573	23	AB018648	Drosophila melanog
35	46.8	6.6	1888	21	AA059321	CDNA encoding a ca
36	46.4	6.6	27541	22	AA017185	Streptomyces nous
37	46.4	6.6	125401	22	AA017186	Streptomyces nous
38	46.2	6.5	642	19	AA023478	Pseudomonas XCP s
39	46.2	6.5	642	21	AA013896	Pseudomonas alcali
40	46.2	6.5	642	21	AA022873	Pseudomonas alcali
41	46.2	6.5	17612	19	AA023494	Pseudomonas XPC, O
42	46.2	6.5	17612	21	AA013905	Pseudomonas alcali
43	46.2	6.5	17612	22	AA030870	Pseudomonas alcali
44	46.2	6.5	17612	24	AA022882	Pseudomonas alcali
45	45.6	6.4	154746	24	AA025519	Human herpesvirus

ALIGNMENTS

RESULT 1
AAF27958
ID AAF27958 standard; DNA; 707 BP.

XX AAF27958;

XX 08-MAY-2001 (first entry)

XX Human dual specificity phosphatase DSP-11 coding sequence.

XX Human dual specificity phosphatase: DSP-11; Duchenne muscular dystrophy;

XX KMP-Kinase signalling cascade; cell proliferation; cancer; allergy;

XX graft-versus-host disease; autoimmune disease; metabolic disease; ss.

XX Homo sapiens.

XX MO200105983-AL.

XX 25-JAN-2001.

XX 19-JUL-2000; 2000MO-US19710.

XX 20-JUL-1999; 99US-0144557.

XX (CEPT-) CEPTYR INC.

XX Luche RM, Wei B;

XX WPI: 2001-147348/15.
XX P-PSDB; AAB35275.
XX Novel dual specificity phosphatase-11 dephosphorylating activated
PT

Applicants

QY 568 GATTTAAACCTTAGCTCCCTGCTGAGACACTGAGTACCCACCTCGACGCA 627
|||||
Db 572 GATTAAACCTTAGCTCCCTGCTGAGACACTGAGTACCCACCTCGACGCA 631
QY 628 GGTCTGATTGAGAGGAGGCTTGTACTGCTTGTGATAAATGATTTACGAACCA 687
|||||
Db 632 GGTCTGATTGAGAGGAGGCTTGTACTGCTTGTGATAAATGATTTACGAACCA 691
QY 688 AAAAAAAAAAAAAAAAAA 707
|||||
Db 692 GGAIAAAAAAAAAAAAAA 711
RESULT 4
ABLA0815
ID ABL40815 standard; cDNA; 746 BP.
XX
AC ABL40815;
XX
DT 03-JUL-2002 (first entry)
XX
DE Human tyrosine phosphatase-like enzyme encoding cDNA.
XX
KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytostatic;
KW cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
KW cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
KW antiarrhythmic; hypotensive; vulnerary; gene therapy; human; gene; ss.
OS Homo sapiens.
XX
PN M0200220747-A2.
XX
PD 14-MAR-2002.
XX
PF 05-SEP-2001; 2001MO-EP10205.
XX
PR 11-SEP-2000; 2000US-231568P.
PR 06-DEC-2000; 2000US-251403P.
XX
PA (FARB) BAYER AG.
XX
PI Kosida S;
XX
DR WPI; 2002-339803/37.
XX
XX
XX New human tyrosine phosphatase-like enzyme polypeptide, regulators of
PT which are useful for preventing, treating diabetes, obesity, cancer,
PT cardiovascular and pulmonary diseases
XX
XX Claim 19; Page 117; 117pp; English.
XX
XX The invention relates to a purified human tyrosine phosphatase-like
CC enzyme polypeptide. The enzyme can be expressed by standard recombinant
CC methodology. The tyrosine phosphatase-like enzyme and encoding
CC polynucleotides are useful for screening for modulators which are used
CC for treating a tyrosine phosphatase-like enzyme dysfunction related
CC disease such as a central nervous system (CNS) disorder, diabetes,
CC obesity, chronic obstructive pulmonary disease, cardiovascular disease,
CC cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
CC other eating disorder such as bulimia. CNS disorders include brain
CC injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
CC Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
CC Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
CC and cardiovascular diseases include myocardial infarction, ischemic
CC diseases of the heart, atrial and ventricular arrhythmia, hypertensive
CC vascular diseases and peripheral vascular diseases. The enzyme is useful
CC in diagnostic assays for detecting diseases and abnormalities or
CC susceptibility to diseases or abnormalities related to the presence of
CC mutations in the encoding nucleic acid sequences. The present sequence
CC represents the human tyrosine phosphatase-like enzyme polypeptide
XX encoding cDNA.
XX

SO Sequence 746 BP; 178 A; 221 C; 218 G; 129 T; 0 other;
Query Match 95.3%; Score 673.6; DB 24; Length 746;
Best Local Similarity 99.4%; Pred. No. 1.8e-131;
Matches 676; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 28 GCGATGGGCGCTGACACCCCACTTCCTGAGTCTTCCGGGCGGCTGGCGAGCTG 87
|||||
Db 32 GCGATGGGCGCTGACACCCCACTTCCTGAGTCTTCCGGGCGGCTGGCGAGCTG 91
QY 88 GCGCTGCGCGGCTCCCGCCACTACAGTTCTGTTGAACTGGGCGTGGCGCACTG 147
|||||
Db 92 GCGCTGCGCGGCTCCCGCCACTACAGTTCTGTTGAACTGGGCGTGGCGCACTG 151
QY 148 GGTGCTGAGGAGGAGGCGGCGGCTCCACAGCGACACTGCGCGGCTACCTGAC 207
|||||
Db 152 GTGTCTCTGAGGAGGAGGCGGCGGCTCCACAGCGACACTGCGCGGCTACCTGAC 211
QY 208 GCGCTGCGCATCCCGCACTTCGCGCGGCGGCGGCGGCGGCGGCGGCTGCGAG 267
|||||
Db 212 GCGCTGCGCATCCCGCACTTCGCGCGGCGGCGGCGGCGGCGGCGGCTGCGAG 271
QY 268 ATCTGAGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 327
|||||
Db 272 ATCTGAGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 331
QY 328 GCGCGCACTGGCACCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 387
|||||
Db 332 GCGCGCACTGGCACCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391
QY 388 GATGCGATTTGCTGAATTCGACGACTACGACCGGCTCATGAGACTATGACGAGAG 447
|||||
Db 392 GATGCGATTTGCTGAATTCGACGACTACGACCGGCTCATGAGACTATGACGAGAG 451
QY 448 AAAGCAGCTTCCAGTTTACACGAGGAAATGAGGCGCTTACCTTACACAG 507
|||||
Db 452 AAAGCAGCTTCCAGTTTACACGAGGAAATGAGGCGCTTACCTTACACAG 511
QY 508 GCGCTGACTCCCTCCCTCCCATGTTGATGGGCGGCAAGTGAAGGAAATGAGCTTAA 567
|||||
Db 512 GCGCTGACTCCCTCCCTCCCATGTTGATGGGCGGCAAGTGAAGGAAATGAGCTTAA 571
QY 568 GATTTAAACCTTAGCTCCCTGCTGAGACACTGAGTACCCACCTCGACGCA 627
|||||
Db 572 GATTTAAACCTTAGCTCCCTGCTGAGACACTGAGTACCCACCTCGACGCA 631
QY 628 GGTCTGATTGAGAGGAGGCTTGTACTGCTTGTGATAAATGATTTACGAACCA 687
|||||
Db 632 GGTCTGATTGAGAGGAGGCTTGTACTGCTTGTGATAAATGATTTACGAACCA 691
QY 688 AAAAAAAAAAAAAAAAAA 707
|||||
Db 692 GGAIAAAAAAAAAAAAAA 711
RESULT 5
ABLA0807
ID ABL40807 standard; DNA; 709 BP.
XX
AC ABL40807;
XX
DT 03-JUL-2002 (first entry)
XX
DE Human tyrosine phosphatase-like enzyme related EST sequence #1.
XX
KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytostatic;
KW cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
KW cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
KW antiarrhythmic; hypotensive; vulnerary; gene therapy; human; gene;
XX expressed sequence tag; EST; ds.
XX
OS Homo sapiens.
XX

PN W0200220747-A2.
 XX 14-MAR-2002.
 PD
 XX 05-SEP-2001; 2001WO-EPI0205.
 PF
 XX 11-SEP-2000; 2000US-231568P.
 PR 06-DEC-2000; 2000US-251403P.
 XX (PARB) BAYER AG.
 PA
 XX Kossida S;
 PI
 DR WPI; 2002-339803/37.
 XX
 PT New human tyrosine phosphatase-like enzyme polypeptide, regulators of
 XX which are useful for preventing, treating diabetes, obesity, cancer,
 PS cardiovascular and pulmonary diseases
 XX
 XX Disclosure; Fig 4; 117pp; English.

XX The invention relates to a purified human tyrosine phosphatase-like
 CC enzyme polypeptide. The enzyme can be expressed by standard recombinant
 CC methodology. The tyrosine phosphatase-like enzyme and encoding
 CC polynucleotides are useful for screening for modulators which are used
 CC for treating a tyrosine phosphatase-like enzyme dysfunction related
 CC disease such as a central nervous system (CNS) disorder, diabetes,
 CC obesity, chronic obstructive pulmonary disease, cardiovascular disease,
 CC cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
 CC other eating disorder such as bulimia. CNS disorders include brain
 CC injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
 CC Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
 CC Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
 CC immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
 CC and cardiovascular diseases include myocardial infarction, ischaemic
 CC diseases of the heart, atrial and ventricular arrhythmia, hypertensive
 CC vascular diseases and peripheral vascular diseases. The enzyme is useful
 CC in diagnostic assays for detecting diseases and abnormalities or
 CC susceptibility to diseases or abnormalities related to the presence of
 CC mutations in the encoding nucleic acid sequences. Sequences AB140807-813
 CC represent human tyrosine phosphatase-like enzyme related expressed
 CC sequence tag (EST) fragments.

XX Sequence 709 BP; 137 A; 233 C; 206 G; 133 T; 0 other;

XX Query Match 95.0%; Score 671.6; DB 24; Length 709;
 XX Best local Similarity 98.6%; Pred. No. 4.7e-131;

XX Matches 688; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 1 TGACCGGCTGCTCTGCTTCCAGCATGGGCGTGCAGCCGCCCACTTCTCTGG 60
 DB 13 TGACCGGCTGCTCTGCTTCCAGCATGGGCGTGCAGCCGCCCACTTCTCTGG 72
 QY 61 GTCCTCCGGGGCGGCGGAGAGTGGCGCTGCGGGCTCCCGGCCACTACCACTTC 120
 DB 73 GTCCTCCGGGGCGGCGGAGAGTGGCGCTGCGGGCTCCCGGCCACTACCACTTC 132
 QY 121 CTCTTGACCTGGGGCGGAGAGTGGTCCCTGACGAGAGGGGGGCCCGCTCACAGC 180
 DB 133 CTCTTGACCTGGGGCGGAGAGTGGTCCCTGACGAGAGGGGGGCCCGCTCACAGC 192
 QY 181 GACAGCTGCCCGGCGCTCACCGCTGACACCGCTCGCATCCCGCACTTCTTCCGCGGCGC 240
 DB 193 GACAGCTGCCCGGCGCTCACCGCTGACACCGCTCGCATCCCGCACTTCTTCCGCGGCGC 252
 QY 241 CCGGACCAATGACCGGCTTCCGAGATGTGGAGAGGCCCAACGACGGGAGAGGCT 300
 DB 253 CCGGACCAATGACCGGCTTCCGAGATGTGGAGAGGCCCAACGACGGGAGAGGCT 312
 QY 301 GGGGAGTGCATGTGCTGGGCTTGGCGGACCTGGGACCAATGCTGGCTTTACTCTG 360
 DB 313 GGGGAGTGCATGTGCTGGGCTTGGCGGACCTGGGACCAATGCTGGCTTTACTCTG 372

QY 361 GTGAAGAGCGGGGCTTGGCTGTCAGAGATGCCATTGCTGTAATCCGACGACTACGACC 420
 DB 373 GTGAAGAGCGGGGCTTGGCTGTCAGAGATGCCATTGCTGTAATCCGACGACTACGACC 432
 QY 421 GGTTCATCGAGACCTTATGAGCAGAGAGAAAGACCTTTCAGTTCTACACGAGAGAA 480
 DB 433 GGTTCATCGAGACCTTATGAGCAGAGAGAAAGACCTTTCAGTTCTACACGAGAGAA 492
 QY 481 TAAAGGCGCTTACACCTTCTACAGAGCCCTCAGCTCCCTCCCAATGTTGATGGG 540
 DB 493 TAAAGGCGCTTACACCTTCTACAGAGCCCTCAGCTCCCTCCCAATGTTGATGGG 552
 QY 541 GCCAGATGAAAGGAGTGAAGTAAAGTAAACCTAGCTCCCATTTGGCTGAGA 600
 DB 553 GCCAGATGAAAGGAGTGAAGTAAAGTAAACCTAGCTCCCATTTGGCTGAGA 612
 QY 601 CACTGAAGTAAAGCCACCCCTGACAGCAGTCCGATTAAGAGGAGGCTTACTGCTT 660
 DB 613 CACTGAAGTAAAGCCACCCCTGACAGCAGTCCGATTAAGAGGAGGCTTACTGCTT 671
 QY 661 GTTGAATTAATGAGTTTACGACCAAAAAAAAAAAAA 698
 DB 672 GTTGAATTAATGAGTTTACGACCAAAAAAAAAAAAA 709

RESULT 6

AAAF63583 standard; cDNA; 694 BP.

AAAF63583;

11-MAY-2001 (first entry)

Human phosphatase BAA91172_h coding sequence.

XX Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
 KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
 KW congenital muscle disorder; Papillon-Lefevre syndrome; Condon disease;
 KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
 KW schizophrenia; hamartoma; ss.

XX Homo sapiens.

XX W0200112819-A2.

XX 22-FEB-2001.

XX 11-AUG-2000; 2000WO-US22158.

XX 13-AUG-1999; 99US-0149005.

XX (SUGEN-) SUGEN INC.

XX Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lloblin M;

XX WPI; 2001-211226/21.

XX P-PsDB; AAB73231.

XX New protein phosphatase polypeptide for diagnosing and treating
 PT phosphatase related disorders such as cancer, schizophrenia, cardiac
 PS dysfunction and/or vascular disorders

XX Claim 1; Fig 4; 138pp; English.

XX The present invention relates to phosphatase proteins and coding
 CC sequences. The present sequence is one such phosphatase coding sequence.
 CC Phosphatases are enzymes that catalyse the dephosphorylation of proteins
 CC modified by phosphorylation of serine, threonine or tyrosine residues.
 CC The phosphatases are useful for treating a variety of diseases: for
 CC example cancer e.g. breast, urogenital, prostate, head, neck, lung
 CC cancers, synovial sarcomas, renal cell carcinoma, non-small cell lung
 CC cancer, hepatocellular carcinoma, pancreatic endocrine tumours, stomach
 CC cancer, glioblastoma, colorectal cancer and thyroid cancer.

CC pathophysiological hypoxia, cardiac dysfunction and/or vascular
CC disorders, myopathies, congenital muscle disorders, papillon-leleuve
CC syndrome, Corden disease, ectodermal dysplasia, Moebius syndrome,
CC Bjornstad syndrome, Bannayan Zonana syndrome, schizophrenia and
CC hamartomas.

Sequence 694 BP; 128 A; 221 C; 216 G; 129 T; 0 other;

Query Match	93.0%	Score 657.4	DB 22	Length 694
Best Local Similarity	99.8%	Pred. No. 4.3e-128		
Matches 658; Conservative	0	Mismatches 1	Indels 0	Gaps 0

QY	28	GCATAGGGGCGTGCAGAGCCGCCCAATTCTCCGCGGGTGTCTCCGGGCCGGCTGGGGGAGCTG	87
Db	32	GGGATGGGCGTGCAGAGCCGCCCAATTCTCCGCGGGTGTCTCCGGGCCGGCTGGGGGAGCTG	91
QY	88	GGCGTGGCGGCGGCTCCCCGGCCCACTACAGTTCCTGTTTGGACCTGGGCGTGGCGGACCTG	147
Db	92	GGCGTGGCGGCGGCTCCCCGGCCCACTACAGTTCCTGTTTGGACCTGGGCGTGGCGGACCTG	151
QY	148	GTGTTCCTGACGGAGCGCGGGGCCCTTCACAGCGACAGCTGCCCGGCTCACCCCTGCAC	207
Db	152	GTGTTCCTGACGGAGCGCGGGGCCCTTCACAGCGACAGCTGCCCGGCTCACCCCTGCAC	211
QY	208	CGCTGGGCAATCCCCGAGCTTGGCCGCGGGCCCCGACACAGCATGACGCTGCCCTTCGTCAG	267
Db	212	CGCTGGGCAATCCCCGAGCTTGGCCGCGGGCCCCGACACAGCATGACGCTGCCCTTCGTCAG	271
QY	268	ATGTGTGAGCAGGCGCAACGCGACGGGGAGAGAGCTGTGGAGTGCACATGTGCTTGGGCTT	327
Db	272	ATGTGTGAGCAGGCGCAACGCGACGGGGAGAGAGCTGTGGAGTGCACATGTGCTTGGGCTT	331
QY	328	GGCGGCACTGGCACATGCTGGGCTGTACTGTGTGAAGAGCGGGGCTTGGCTGCAGGA	387
Db	332	GGCGGCACTGGCACATGCTGGGCTGTACTGTGTGAAGAGCGGGGCTTGGCTGCAGGA	391
QY	388	GATGCATTTGCTGAATTCGACGACACTGACGACCCGGCTCATTCGAGACCTTATGACAGAG	447
Db	392	GATGCATTTGCTGAATTCGACGACACTGACGACCCGGGCTCATTCGAGACCTTATGACAGAG	451
QY	448	AAAGCACTCTCCAGTCTCTACAGACGGAAGAAATAAAGGGCTTTAGTACCCTTCTACAG	507
Db	452	AAAGCACTCTCCAGTCTCTACAGACGGAAGAAATAAAGGGGCTTTAGTACCCTTCTACAG	511
QY	508	GGCCTCACTCCCTTCCCCAGATGTGTGATGAGGGGCGAGATGAGAAGGAAGTGAAGCTAAA	567
Db	512	GGCCTCACTCCCTTCCCCAGATGTGTGATGAGGGGCGAGATGAGAAGGAAGTGAAGCTAAA	571
QY	568	GTAATTAACCTCTAGCTCCCATTTGGCTGAAGACACTGAAGTAGCCACCCCTGCAGGCA	627
Db	572	GTAATTAACCTCTAGCTCCCATTTGGCTGAAGACACTGAAGTAGCCACCCCTGCAGGCA	631
QY	628	GGTCCGTATGGAAGGGAGGCTGTACTGCTTTGTGAATAAATGATTTACGAGCA	686
Db	632	GGTCCGTATGGAAGGGAGGCTGTACTGCTTTGTGAATAAATGATTTACGAGCA	690

RESULT 7
ABL40810
ID ABL40810 standard; DNA; 602 BP.

AC ABL40810

DT 03-JUL-2002 (first entry)

Human tyrosine phosphatase-like enzyme related EST sequence #4

KM Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytosolatic;
KM cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant
KM cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
KM antianxietytic; hypotensive; vulnerrary; gene therapy; human; gene;
KM expressed sequence tag; EST; ds.

OS Homo sapiens.

PN WO200220747-A2

PD 14-MAR-2002

PF 05-SEP-2001; 2001WO-EP10205.
XX

PR 11-SEP-2000; 2000US-231568P.
DE 06-DEC-2000; 2000US-231403D

XX
XX
XX (EARTH) DAYTON MOXX
XX
PT

KONGSBERG

XX
DP
FBI, 2003-

XX
XX
New buses and trucks

PT New human tyrosine phosphatase-like enzyme polypeptide, regulators of
PT which are useful for preventing, treating diabetes, obesity, cancer,
PT cardiovascular and pulmonary diseases -

PS Disclosure; Fig 7; 117pp; English.

The invention relates to a purified human tyrosine phosphatase-like enzyme polypeptide. The enzyme can be expressed by standard recombinant methodology. The tyrosine phosphatase-like enzyme and encoding polynucleotides are useful for screening for modulators which are used for treating a tyrosine phosphatase-like enzyme dysfunction related disease such as a central nervous system (CNS) disorder, diabetes, obesity, chronic obstructive pulmonary disease, cardiovascular disease, cancer, anorexia, cachexia, wasting disorders, appetite suppression, and other eating disorder such as bulimia. CNS disorders include brain injuries, Parkinson's disease, dementia, multiple sclerosis, stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human immunodeficiency virus (HIV) dementia, pain associated with CNS disorders and cardiovascular diseases include myocardial infarction, ischemic diseases of the heart, atrial and ventricular arrhythmia, hypertensive vascular diseases and peripheral vascular diseases. The enzyme is useful in diagnostic assays for detecting diseases and abnormalities or susceptibility to diseases or abnormalities related to the presence of mutations in the encoding nucleic acid sequences. Sequences AB140807-813 represent human tyrosine phosphatase-like enzyme related expressed sequence tag (EST) fragments.

SQ Sequence 602 BP; 109 A; 202 C; 179 G; 110 T; 2 other,

Query Match	82.6%	Score 583.8	DB 24	Length 602
Best Local Similarity	99.3%	Pred. No. 9.5e-113		
Matches 596	Conservative	0	Mismatches 13	Indels 1
				Gaps 1

QY	42	GC	CCCCAACCTTCTCCCTGGGGGCTCCGGGCGGGCTGGGGAGACTGGCGCTGGCGGGCT	101
Db	3	GG	CGCCCAACTTCTCTCGGGGTGCTCCGGGCGGGCTGGGGAGACTGGCGCTGGCGGGCT	62
QY	102	CC	CGCGCCCACTACCAAGTTCTGTGTGACACTGGGCGTTCGGCACTGGTGTCCCTGACGA	161
Db	63	CC	CGCGCCCACTACCAAGTTCTGTGTGACACTGGGCGTTCGGCACTGGTGTCCCTGACGA	122
QY	162	GC	GGGGGGGCCCCCTACACAGCGACAGCTGCCCGGCTACCTTCACACCGCGCTGGCATCCC	221
Db	123	GC	GGGGGGGCCCCCTACACAGCGACAGCTGCCCGGCTACCTTCACACCGCGCTGGCATCCC	182
QY	222	CG	AATTGTGCGCCGGCGGCGCCCGACAGATGACACCGTTCTGTGCAGATGCTGGACAGGC	281
Db	183	CG	AATTGTGCGCCGGCGGCGCCCGACAGATGACACCGTTCTGTGCAGATGCTGGACAGGC	242
QY	282	CA	ACGACGACGGGAGAGAGGCTGTGGGAGTGCACCTATGCTTGGGCTTGGCGGCACTGGCAC	341
Db	243	CA	ACGACGACGGGAGAGAGGCTGTGGGAGTGCACCTATGCTTGGGCTTGGCGGCACTGGCAC	302
QY	342	CA	TGCTGCGCTGTATTACCTGGTGA- GGAGCGGGGCTTGGCTGACGAGATGCAATTGCTT	400

Db 303 CAGCTGCGCTGTACTGCTGAAGGAGCGGNCCTTGGCTGACGAGATGCCATTGCTG 362
 Qy 401 AATCCGACGACCTACGACCGGCTCATCGACCTTATGACGAGAAAGCACTCTCC 460
 Db 363 AATTCGACGACCTACGACCGGCTCATCGACCTTATGACGAGAAAGCACTCTCC 422
 Qy 461 AGTTCTACGACGAGAAATTAAGGGGCTTATGACCTTCTACGAGGCCCTCACTCC 520
 Db 423 AGTTCTACGACGAGAAATTAAGGGGCTTATGACCTTCTACGAGGCCCTCACTCC 482
 Qy 521 TTCCCTCATGTTGTCGATGGGGCCGAGATGAAGGAAATGAGTAAATTAACCTC 580
 Db 483 TTCCCTCATGTTGTCGATGGGGCCGAGATGAAGGAAATGAGTAAATTAACCTC 542
 Qy 581 TACCTCCATTGGCTGAAGACACTGAAGTATGCCACCCCTGACGAGGATCGATGAA 640
 Db 543 TACCTCCATTGGCTGAAGACACTGAAGTATGCCACCCCTGACGAGGATCGATGAA 602
 RESULT 8
 ID ABL40812 standard; DNA: 553 BP.
 XX ABL40812:
 AC
 DT 03-JUL-2002 (first entry)
 XX
 DE Human tyrosine phosphatase-like enzyme related EST sequence #6.
 XX
 KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cyostatic;
 KW cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
 KW cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
 KW antiarhythmic; hypotensive; vulnerary; gene therapy; human; gene;
 KW expressed sequence tag; EST; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200220747-A2.
 PD 14-MAR-2002.
 XX
 PF 05-SEP-2001; 2001WO-EP10205.
 XX
 PR 11-SEP-2000; 2000US-231568P.
 XX 06-DEC-2000; 2000US-251403P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Kosiida S;
 XX
 DR WPI; 2002-339803/37.
 XX
 PT New human tyrosine phosphatase-like enzyme polypeptide, regulators of
 PT which are useful for preventing, treating diabetes, obesity, cancer,
 PT cardiovascular and pulmonary diseases
 XX
 PS Disclosure; Fig 9; 117pp; English.
 XX
 CC The invention relates to a purified human tyrosine phosphatase-like
 CC enzyme polypeptide. The enzyme can be expressed by standard recombinant
 CC methodology. The tyrosine phosphatase-like enzyme and encoding
 CC polynucleotides are useful for screening for modulators which are used
 CC for treating a tyrosine phosphatase-like enzyme dysfunction related
 CC disease such as a central nervous system (CNS) disorder, diabetes,
 CC obesity, chronic obstructive pulmonary disease, cardiovascular disease,
 CC cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
 CC other eating disorder such as bulimia. CNS disorders include brain
 CC injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
 CC Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
 CC Creutzfeldt-Jacob disease, progressive nuclear palsy, and human
 CC immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
 CC and cardiovascular diseases include myocardial infarction, ischaemic
 CC diseases of the heart, atrial and ventricular arrhythmia, hypertensive

CC vascular diseases and peripheral vascular diseases. The enzyme is useful
 CC in diagnostic assays for detecting diseases and abnormalities or
 CC susceptibility to diseases or abnormalities related to the presence of
 CC mutations in the encoding nucleic acid sequences. Sequences ABL40807-813
 CC represent human tyrosine phosphatase-like enzyme related expressed
 CC sequence tag (EST) fragments.
 XX
 SQ Sequence 553 BP; 96 A; 186 C; 168 G; 103 T; 0 other;
 Query Match 71.9%; Score 508.6; DB 24; Length 553;
 Best Local Similarity 98.6%; Pred. No. 4, 6e-97;
 Matches 545; Conservative 0; Mismatches 4; Indels 4; Gaps 3;
 Qy 29 CGATGGGCGTGGACGCCCCCACTTCCTCCGCGCTCTCCGGCCGCTGGCGGACTGG 88
 Db 1 CGATGGGCGTGGACGCCCCCACTTCCTCCGCGCTCTCCGGCCGCTGGCGGACTGG 60
 Qy 89 CGCTGCGCGGCTCCCGCCACACACAGTTCCTGTTGACACCTGGCGGCTGGCACTGG 148
 Db 61 CGCTGCGCGGCTCCCGCCACACACAGTTCCTGTTGACACCTGGCGGCTGGCACTGG 120
 Qy 149 TGTCCCTGAC--GGAGCGGCGGCCCCCTCACAGGACAGTGGCCGGCTCACCTGCA 206
 Db 121 TGTCCCTGACAGTGGACGCTGGCGGCCCCCTCACAGGACAGTGGCCGGCTCACCTGCA 180
 Qy 207 CCGCTGCGCATGCC--GACTTCGCGCGCGGCCCCCGACAGATGACCGGCTGGTGC 265
 Db 181 CCGCTGCGCATGCCCGCTGACTTCTGCCCCGCGCCCCCGACAGATGACCGGCTGGTGC 240
 Qy 266 AGATCGTGAAGA--GGCCAAAGCAGCGGGAGAGGCTGGGAGTGAAGTGTCTGGGCG 324
 Db 241 AGATCGTGAAGAAGTGGCCAAAGCAGCGGGAGAGGCTGGGAGTGAAGTGTCTGGGCG 300
 Qy 325 TTGGGCGGACACTGGACATGCTGGCTGTACTGTGTAAGAGCGGGCTTGGCTGCA 384
 Db 301 TTGGGCGGACACTGGACATGCTGGCTGTACTGTGTAAGAGCGGGCTTGGCTGCA 360
 Qy 385 GGAGATGCCATTTGCTGAATCCGACGACTACGACCGGCTCCATGAGACTTATGAGCAG 444
 Db 361 GGAGATGCCATTTGCTGAATCCGACGACTACGACCGGCTCCATGAGACTTATGAGCAG 420
 Qy 445 GAGAAAGCAGTCTTCAGTTCTACGAGCAGAAATTAAGGGGCTTATGACCTTCTAC 504
 Db 421 GAGAAAGCAGTCTTCAGTTCTACGAGCAGAAATTAAGGGGCTTATGACCTTCTAC 480
 Qy 505 CAGGCCCTCACTCCCTTCCCATGTTGTGATGGGGCCAGAGTGAAGGAAAGTGACT 564
 Db 481 CAGGCCCTCACTCCCTTCCCATGTTGTGATGGGGCCAGAGTGAAGGAAAGTGACT 540
 Qy 565 AAAGTATTAAC 577
 Db 541 AAAGTATTAAC 553
 RESULT 9
 ID ABL40808 standard; DNA: 493 BP.
 XX ABL40808:
 AC
 DT 03-JUL-2002 (first entry)
 XX
 DE Human tyrosine phosphatase-like enzyme related EST sequence #2.
 XX
 KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cyostatic;
 KW cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
 KW cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
 KW antiarhythmic; hypotensive; vulnerary; gene therapy; human; gene;
 KW expressed sequence tag; EST; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200220747-A2.

CC blood-related disorders (e.g., anemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
S0 Sequence 559 BP; 90 A; 192 C; 169 G; 101 T; 7 other;

Query Match 66.7%; Score 471.8; DB 24; Length 559;
Best Local Similarity 99.6%; Pred. No. 2.2e-89;
Matches 473; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 15 GTGCGCTTCCAGGATGGGCGTGCAGCCGCCCAACTTCCTGGGCTTCGCGGCG 74
DB 72 GTCCGCTTCCAGGATGGGCGTGCAGCCGCCCAACTTCCTGGGCTTCGCGGCG 131
OY 75 GCTGGCGGAGCTGGCGCTGCGCGGCTCCCGCCCACTACAGATTCTGTTGAGACTGG 134
DB 132 GCTGGCGGAGCTGGCGGCTGCGCGGCTCCCGCCCACTACAGATTCTGTTGAGACTGG 191
OY 135 GCTGGCGGAGCTGGCTGTCTCTGAGAGAGCGGGCGCCCTACAGAGAGCTGCCGCG 194
DB 192 GCTGGCGGAGCTGGCTGTCTCTGAGAGAGCGGGCGCCCTACAGAGAGCTGCCGCG 251
OY 195 CCTACCCCTGACCGCGCTGGCGATCCCGACTTCCTGGCGGCGCCCGCCAGCAATGA 254
DB 252 CCTACCCCTGACCGCGCTGGCGATCCCGACTTCCTGGCGGCGCCCGCCAGCAATGA 311
OY 255 CCGCTTCTGAGATCGTGAAGAGAGCCAGCAAGGAGAGGCTGTGGAGTGCATG 314
DB 312 CCGCTTCTGAGATCGTGAAGAGAGCCAGCAAGGAGAGGCTGTGGAGTGCATG 371
OY 315 TGGCTTGGGCTTGGCGGCACTGGGACCATGTGCTTACTCTGTGGAAGAGCGGG 374
DB 372 TGGCTTGGGCTTGGCGGCACTGGGACCATGTGCTTACTCTGTGGAAGAGCGGG 431
OY 375 CTGAGCTGACAGAGATGCTTGAATCGAGAGTACAGACCGGCTCCATGAGAGC 434
DB 432 CTGAGCTGACAGAGATGCTTGAATCGAGAGTACAGACCGGCTCCATGAGAGC 491
OY 435 CTATGAGCAGAGAAAGAGTCTTCACTTACAGAGAAAGAAATGAAGGCC 489
DB 492 CTATGAGCAGAGAAAGAGTCTTCACTTACAGAGAAAGAAATGAAGGCC 546

RESULT 11
ABL40813/c
ID ABL40813 standard; DNA; 448 BP.
XX
AC ABL40813;
XX
DT 03-JUL-2002 (first entry)
XX
XX Human tyrosine phosphatase-like enzyme related EST sequence #7.
XX
KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytosolic;
KW cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
KW cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
KW antidiabetic; hypotensive; vulnary; gene therapy; human; gene;
XX expressed sequence tag; EST; ds.
XX
XX Homo sapiens.
XX
PN WO200220747-A2.

XX
PD 14-MAR-2002.
XX
XX 05-SEP-2001; 2001WO-EP10205.
XX
XX 11-SEP-2000; 2000US-231568P.
XX 06-DEC-2000; 2000US-251403P.
XX
XX (FARB) BAYER AG.
XX
XX Kossida S;
XX
XX WPI; 2002-339803/37.
XX
XX New human tyrosine phosphatase-like enzyme polypeptide, regulators of
XX which are useful for preventing, treating diabetes, obesity, cancer,
XX cardiovascular and pulmonary diseases
XX
XX Disclosure; Fig 10; 117p; English.
XX
XX The invention relates to a purified human tyrosine phosphatase-like
XX enzyme polypeptide. The enzyme can be expressed by standard recombinant
XX methodology. The tyrosine phosphatase-like enzyme and encoding
XX polynucleotides are useful for screening for modulators which are used
XX for treating a tyrosine phosphatase-like enzyme dysfunction related
XX disease such as a central nervous system (CNS) disorder, diabetes,
XX obesity, chronic obstructive pulmonary disease, cardiovascular disease,
XX cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
XX other eating disorder such as bulimia. CNS disorders include brain
XX injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
XX Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
XX Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
XX immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
XX and cardiovascular diseases include myocardial infarction, ischemic
XX diseases of the heart, atrial and ventricular arrhythmia, hypertensive
XX vascular diseases and peripheral vascular diseases. The enzyme is useful
XX in diagnostic assays for detecting diseases and abnormalities or
XX susceptibility to diseases or abnormalities related to the presence of
XX mutations in the encoding nucleic acid sequences. Sequences ABL40807-813
XX represent human tyrosine phosphatase-like enzyme related expressed
XX sequence tag (EST) fragments.
XX
S0 Sequence 448 BP; 93 A; 132 C; 120 G; 103 T; 0 other;

Query Match 63.1%; Score 446; DB 24; Length 448;
Best Local Similarity 100.0%; Pred. NO. 5.1e-84;
Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 241 CCGACGATCGACCGCTTGTGAGATCGTGAAGAGCCAGCAAGGAGAGGCT 300
DB 448 CCGACGATCGACCGCTTGTGAGATCGTGAAGAGCCAGCAAGGAGAGGCT 389
OY 301 GTGGAGTGCATGTGCTGTGGCTTGGCGGCACTGGCACCACCATGCTGCTTACCTG 360
DB 388 GTGGAGTGCATGTGCTGTGGCTTGGCGGCACTGGCACCACCATGCTGCTTACCTG 329
OY 361 GTGAAGAGCGGGGCTTGGCTGAGAGATCGATGCTGAATCCGACGACTACGACC 420
DB 328 GTGAAGAGCGGGGCTTGGCTGAGAGATCGATGCTGAATCCGACGACTACGACC 269
OY 421 GGCTCCATCGAGACCTATGAGCAGAGAAACAGTCTTCCAGTTCTACAGAGAA 480
DB 268 GGCTCCATCGAGACCTATGAGCAGAGAAACAGTCTTCCAGTTCTACAGAGAA 209
OY 481 TAAAGGGGCTTAAAGTACCTTCTTACAGAGCCCTCACTCCCTCCAGTGTGATGG 540
DB 208 TAAAGGGGCTTAAAGTACCTTCTTACAGAGCCCTCACTCCCTCCAGTGTGATGG 149
OY 541 GCCAGAGATGAAGGAATGAGCTAAATATTAACCTTACGTCCTCCATGGCTGAAGA 600
DB 148 GCCAGAGATGAAGGAATGAGCTAAATATTAACCTTACGTCCTCCATGGCTGAAGA 89
OY 601 CACTGAAGTACGCCACCCCTGACAGGAGTCTGATTGAAGGAGGCTGTACTGCTT 660

```

|||||
DB 88 CACTGATGAGCCACCCCTGCAGGAGGTCTGATTGAAGGAGGCTGTGCTCTT 29
QY 661 GTTGAATTAATGAGTTTACGAACA 686
DB 28 GTTGAATTAATGAGTTTACGAACA 3

RESULT 12
ABLA0809
ID ABLA0809 standard; DNA; 453 BP.
XX
AC ABLA0809;
XX
DE 03-JUL-2002 (first entry)
XX
DE Human tyrosine phosphatase-like enzyme related EST sequence #3.
XX
KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytosolic;
KW cardiac; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
KW cerebrotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
KW antiarrhythmic; hypotensive; vulnerary; gene therapy; human; gene;
KW expressed sequence tag; EST; ds.
XX
OS Homo sapiens.
XX
PN WO200220747-A2.
XX
PD 14-MAR-2002.
XX
PF 05-SEP-2001; 2001WO-EP10205.
XX
PR 11-SEP-2000; 2000US-231568P.
PR 06-DEC-2000; 2000US-251403P.
XX
PA (FARB ) BAYER AG.
PI Kossida S;
XX
DR WPI: 2002-339803/37.
XX
PT New human tyrosine phosphatase-like enzyme polypeptide, regulators of
PT which are useful for preventing, treating diabetes, obesity, cancer,
PT cardiovascular and pulmonary diseases
XX
PS Disclosure: Fig 6; 117pp; English.
XX
CC The invention relates to a purified human tyrosine phosphatase-like
CC enzyme polypeptide. The enzyme can be expressed by standard recombinant
CC methodology. The tyrosine phosphatase-like enzyme and encoding
CC polynucleotides are useful for screening for modulators which are used
CC for treating a tyrosine phosphatase-like enzyme dysfunction related
CC disease such as a central nervous system (CNS) disorder, diabetes,
CC obesity, chronic obstructive pulmonary disease, cardiovascular disease,
CC cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
CC other eating disorder such as bulimia. CNS disorders include brain
CC injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
CC Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
CC Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
CC and cardiovascular diseases include myocardial infarction, ischaemic
CC diseases of the heart, atrial and ventricular arrhythmia, hypertensive
CC vascular diseases and peripheral vascular diseases. The enzyme is useful
CC in diagnostic assays for detecting diseases and abnormalities or
CC susceptibility to diseases or abnormalities related to the presence of
CC mutations in the encoding nucleic acid sequences. Sequences ABL40807-813
CC represent human tyrosine phosphatase-like enzyme related expressed
CC sequence tag (EST) fragments.
XX
SQ Sequence 453 BP; 67 A; 161 C; 153 G; 72 T; 0 other;
Query Match 60.7%; Score 429; DB 24; Length 453;
Best Local Similarity 100.0%; Pred. No. 1.8e-80;

```

```

Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 CCAGCGATGGGGCTGACGCCCCCAACTTCTCTGAGTCTTCCGGCCCGCTGGCGGA 84
DB 25 CCAGCGATGGGGCTGACGCCCCCAACTTCTCTGAGTCTTCCGGCCCGCTGGCGGA 84
QY 85 CTGGGCTGCGCGGCTCCCGCCCACTACAGTTCTGTGACCTGGGCTGGGGC 144
DB 85 CTGGGCTGCGCGGCTCCCGCCCACTACAGTTCTGTGACCTGGGCTGGGGC 144
QY 145 CTGGTGTCCCTGACGAGGAGCGGGGCCCTCAGAGGACAGTGGCCCGGCTGACCCG 204
DB 145 CTGGTGTCCCTGACGAGGAGCGGGGCCCTCAGAGGACAGTGGCCCGGCTGACCCG 204
QY 205 CACCGCTGCGGATCCCGCACTTCTGCCCCCGCCCGCCGACAGATCGACGCTTCTG 264
DB 205 CACCGCTGCGGATCCCGCACTTCTGCCCCCGCCCGCCGACAGATCGACGCTTCTG 264
QY 265 CAGATCTGTGACGAGGCCAACGACGAGGAGAGGCTGTGGAGTGCATGTCTGTGGC 324
DB 265 CAGATCTGTGACGAGGCCAACGACGAGGAGAGGCTGTGGAGTGCATGTCTGTGGC 324
QY 325 TTTGGCGGCACTGGGACCATGCTGGCCGTACTGCTGTAAGGACGCGGCTTGGCTGA 384
DB 325 TTTGGCGGCACTGGGACCATGCTGGCCGTACTGCTGTAAGGACGCGGCTTGGCTGA 384
QY 385 GGAGATGCCATTCTGAATCCGACGACATACGCCGCTCCATGAGACATATGAGCAG 444
DB 385 GGAGATGCCATTCTGAATCCGACGACATACGCCGCTCCATGAGACATATGAGCAG 444
QY 445 GAGAAAGCA 453
DB 445 GAGAAAGCA 453

```

```

RESULT 13
AAF63563
ID AAF63563 standard; cDNA; 715 BP.
XX
AC AAF63563;
XX
DE 11-MAY-2001 (first entry)
XX
DE Murine phosphatase AA023073_m coding sequence.
XX
KW phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
KW congenital muscle disorder; Papillon-Léve syndrome; Cowden disease;
KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
KW schizophrenia; Hamman-Rich syndrome; ss.
XX
OS Mus sp.
XX
PN WO200112819-A2.
XX
PD 22-FEB-2001.
XX
PF 11-AUG-2000; 2000WO-US22158.
XX
PR 13-AUG-1999; 99US-0149005.
XX
PA (SUGEN-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Llaoulin M;
XX
DR WPI: 2001-211226/21.
XX
DR P-PSDB; AAB73211.
XX
PT New protein phosphatase polypeptide for diagnosing and treating
PT phosphatase related disorders such as cancer, schizophrenia, cardiac
PT dysfunction and/or vascular disorders
XX
PS Claim 1; Fig 4; 138pp; English.

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: November 4, 2002, 08:49:32 ; Search time 48 Seconds
(without alignments)
4907.093 Million cell updates/sec

Title: US-09-619-380-1

Perfect score: 707
Sequence: 1 tgcacgcgtcctcgtcc.....aaaaaaaaaaaaaaa 707

Scoring table: IDENTITY_NNC
Gapop 10.0, Gapext 1.0

Searched: 310279 seqs, 166577418 residues

Total number of hits satisfying chosen parameters: 620558

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	241.8	34.2	247	10	US-09-964-824A-73
2	49.2	7.0	1169	10	US-09-747-835A-11
3	49.2	7.0	1467	10	US-09-747-835A-14
4	49.2	7.0	2936	10	US-09-747-835A-12
5	45	6.4	3721	12	US-10-036-342-34
6	44.6	6.3	1366	10	US-09-917-800A-1527
7	44.6	6.3	2455	10	US-09-917-800A-1386
8	44.4	6.3	1680	10	US-09-815-242-7988
9	43.8	6.2	1815	10	US-09-758-269-13
C 10	43.6	6.2	4282	10	US-09-778-927A-16
C 11	43.6	6.2	4366	12	US-10-044-090-668
12	42.8	6.1	6000	10	US-09-764-860-768
13	42.2	6.0	1347	10	US-09-747-835A-10
C 14	42.2	6.0	1347	10	US-09-815-242-7912
15	42	5.9	298	10	US-09-759-143-399
16	42	5.9	298	10	US-09-780-666-399
17	42	5.9	298	10	US-09-822-827-399
18	42	5.9	350	10	US-09-925-301-677
19	42	5.9	2419	10	US-09-939-408A-12

20	42	5.9	2741	10	US-09-939-408A-28	Sequence 28, Appl
21	41.6	5.9	891	10	US-09-739-907-25	Sequence 25, Appl
22	41.4	5.9	2636	12	US-10-002-600-95	Sequence 95, Appl
23	41.4	5.9	15872	10	US-09-861-289-1	Sequence 1, Appl
24	41.2	5.8	8195	10	US-09-795-693-7	Sequence 7, Appl
25	41	5.8	1929	10	US-09-761-534A-9	Sequence 9, Appl
26	40.4	5.7	2658	10	US-09-815-242-4035	Sequence 4035, Ap
27	40.4	5.7	11220	10	US-09-861-289-32	Sequence 32, Ap
28	40.4	5.7	36778	10	US-09-861-289-5	Sequence 5, Appl
29	40.2	5.7	500	10	US-09-754-997A-43	Sequence 43, Appl
30	40.2	5.7	800	10	US-09-866-987-2	Sequence 2, Appl
31	40.2	5.7	821	10	US-09-840-787-70	Sequence 70, Appl
32	40.2	5.7	2382	10	US-09-789-404-3	Sequence 3, Appl
C 33	40.2	5.7	2636	10	US-09-889-404-1	Sequence 1, Appl
C 34	40.2	5.7	3350	10	US-09-894-998-48	Sequence 48, Appl
C 35	40.2	5.7	14800	10	US-09-954-456-1601	Sequence 1601, Ap
C 36	40	5.7	1644	10	US-09-815-242-7884	Sequence 7884, Ap
C 37	39.8	5.6	763	10	US-09-919-580-871	Sequence 871, Ap
38	39.8	5.6	1402	9	US-09-931-457A-44	Sequence 44, Appl
39	39.6	5.6	2125	10	US-09-880-107-2412	Sequence 2412, Ap
C 40	39.6	5.6	2280	10	US-09-815-242-4005	Sequence 4005, Ap
41	39.6	5.6	3955	10	US-09-855-722-4	Sequence 4, Appl
42	39.6	5.6	4702	10	US-09-880-107-1545	Sequence 1545, Ap
43	39.4	5.6	1138	10	US-09-778-844-165	Sequence 165, App
44	39.4	5.6	1157	10	US-09-778-844-165	Sequence 165, App
45	39.4	5.6	1183	10	US-09-778-844-167	Sequence 167, App

ALIGNMENTS

RESULT 1
US-09-964-824A-73/c
; Sequence 73, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sigr
; FILE REFERENCE: 869290-73
; CURRENT APPLICATION NUMBER: US/09/964, 824A
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236, 033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 028
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73
; LENGTH: 247
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-73
Query Match 34.2% Score 241.8; DB 10; Length 247;
Best Local Similarity 99.2% Pred. No. 1.2e-45;
Matches 245; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CAGAGAAACAGAGCTCCAGCTTACACAGAGAAATTAAGGGGCTTAGTACCCCTTC 501
|||||
CAGAGAAACAGAGCTCCAGCTTACACAGAGAAATTAAGGGGCTTAGTACCCCTTC 188
|||||
TACACAGCCCTCAGCTCCCTCCATGTTGTCAGTGGGCGACAGATGAAGGAATGCG 128
|||||
ACATAATATTAACCTCTAGCTCCCATGCTGGAAGACACTGAATAGCCACCCCTG 621
|||||
ACATAATATTAACCTCTAGCTCCCATGCTGGAAGACACTGAATAGCCACCCCTG 68
|||||
CAGCAGGCTCCTGATGAAGGAGGAGCTTACTGCTTGTGTAATTAAGAGTTTACG 681

```
Db 67 CAGCAGGCTGCTGATTAAGAGGAGGCTTGACTGCTTTGATGATTAATGATTTACG 8
OY 682 AACCA 686
7 AACCA 3

RESULT 2
US-09-747-835A-11
; Sequence 11, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrul
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Dmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; FILE REFERENCE: HYS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747, 835A
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 1169
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-747-835A-11

Query Match
Best Local Similarity 46.7%; Score 49.2; DB 10; Length 1169;
Matches 156; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

OY 65 TTCCGGGCGGCTGGGAGACTGGCGCTGCCGCGCTCCCGCCACTACCAAGTTCTGT 124
Db 109 TGCCCGGCGTGTGGCTGGCTGCTGCAAGTGGCGGCGGACCAAGTACAGGGCTTCC 168
OY 125 TGAACCTGGGCGTGGCGACCTGCTGCTGCTGCAAGGAGCGGGGCCCCCTACAGCGACA 184
Db 169 CCGAGCTGGCTGGAGCACTGGCTTACAGACCGCAAGATCGGGCTCTCAGCTTCTTCT 228
OY 185 GCTGCCCCGCGCTACCTGCAACCGCTGGCGATGCCGACTTGTGCGCGGCGCCCGC 244
Db 229 GCGCGCGCTGCAAGCGCTTCTACAGCTTGTGCTTGGCGGCGGCGCCACCGCTACG 288
OY 245 ACCAGATGACCGCTTGTGCAAGATCGTGAAGGCAAGCAAGCGGAGAGGCTGTGG 304
Db 289 ACCTGGTCAACCTGGAGTCAAGAGAGTCTTGGGCAACAGAGCACTCTGGGTGAGG 348
OY 305 GAGTGCACTGTGCTGCTGGCTTGGCGGCACTGGCGACCAATGCTGTACCTGTGA 364
Db 349 AGGAGGCTGTGGCGGATGAGATCTACCTTCCCTGGGAGTGTGCGCTCGGACAGTTGT 408
OY 365 AGGAGCGGGGCTGGCTGCAAGAGATGCCAATTGC 398
```

```
Db 409 CCTGCTGGCGGCTGACTGACTGCCGCTCAATTGC 442

RESULT 3
US-09-747-835A-14
; Sequence 14, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrul
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Dmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-L
; FILE REFERENCE: HYS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747, 835A
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-747-835A-14

Query Match
Best Local Similarity 46.7%; Score 49.2; DB 10; Length 1467;
Matches 156; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

OY 65 TTCCGGGCGGCTGGGAGACTGGCGCTGCCGCGGCTCCCGCCACTACCAAGTTCTGT 124
Db 815 TGCCCGGCGTGTGGCGGCTGCTGCAAGTGGCGGCGGACCAAGTACAGGCTTCC 874
OY 125 TGAACCTGGGCGTGGCGACCTGCTGCTGCTGCAAGGAGCGGGGCCCCCTACAGCGACA 184
Db 875 CCGAGCTGGCTGGAGCACTGGCTTACAGACCGCAAGATGAGGCTGCTCAGCTTCTTCT 934
OY 185 GCTGCCCCGCGCTACCTGCAACCGCTGGCGATGCCGACTTGTGCGCGGCGCCCGC 244
Db 935 GCGCGCGCTGCAAGCGCTTACAGCTTGTGCTGCGGCTGCGCGGCGCCACCGCTACG 994
OY 245 ACCAGATGACCGCTTGTGCAAGATCGTGAAGGCAAGGCAAGCAAGCGGAGAGGCTGTG 304
Db 995 ACCTGGTCAACCTGGAGTCAAGAGTCTTGGGCAACAGAGCACTCTGGGTGAGG 1054
OY 305 GAGTGCACTGTGCTGCTGGCTTGGCGGCACTGGCGACCAATGCTGTACCTGTGA 364
Db 1055 AGGAGGCTGTGGCGGATGAGATCTACCTTCCCTGGGAGTGTGCGCTCGGACGTTGT 1114
OY 365 AGGAGCGGGGCTGGCTGCAAGAGATGCCAATTGC 398
Db 1115 CCTGCTGGCGGCTGACTGACTGCCGCTCAATTGC 1148

RESULT 4
US-09-747-835A-12
```

```
Sequence 12, Application US/09/747,835A
Patent No. US20020146692A1
GENERAL INFORMATION:
APPLICANT: Yamazaki, Victoria
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Zhou, Ping
APPLICANT: Wang, Dunru
APPLICANT: Zhang, Jie
APPLICANT: Ren, Felyan
APPLICANT: Asundi, Vinod
APPLICANT: Dumanac, Radoje T
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
FILE REFERENCE: HYS-37CIP
CURRENT APPLICATION NUMBER: US/09/747,835A
PRIOR FILING DATE: 2002-03-08
PRIOR APPLICATION NUMBER: US 09/729,739
PRIOR FILING DATE: 2000-12-04
PRIOR APPLICATION NUMBER: US 09/653,450
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: US 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US 09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: US 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 63
SOFTWARE: PatentIn version 3.0
SEQ ID NO 12
LENGTH: 2936
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (135)..(1601)
US-09-747-835A-12
```

```
Query Match          7.0%; Score 49.2; DB 10; Length 2936;
Best Local Similarity 46.7%; Pred. No. 0.028;
Matches 156; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 65 TTCCGGGCGGCTGGCGGAGCTGGCGCCGCTCCCGCCACTACAGTTCTGT 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 949 TGGCGGCGGCTGGCGGCGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 125 TGGACCTGGGCGTGGCGGACCTGGTGTCCCTGACGGAGCGGCGGCGGCGG 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1009 CCGACTGGCTGGACACCACTGGCTACAGCAGCGGAGATCGGGCTGCTCAGCTTCT 1068
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 185 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1069 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 245 ACCAGATGACGCGCTTGGCAGATGCTGGAGCAGGCGGCGGAGAGAGCTGG 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1129 ACTGTGTAACCTGGAGGTAACAGAGTCTGGGCGGCGGCGGCGGCGGCGG 1188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 305 GAGTGCATCTGTCTGGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1189 AGAGGTCTGGCGGAGATGATCTACCTCCCTGGGAGATGCGGCGGCGGCG 1248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 365 AGAGCGGCGGCTGGCTGAGAGGAGATGCCATTGC 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1249 CCTGTGGCGGCTGACCTACTGCGCTCATTTGC 1282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 5
US-10-036-342-34
Sequence 34, Application US/10036342
Patent No. US20020090681A1
```

```
GENERAL INFORMATION:
APPLICANT: Desnoyers, Luc
APPLICANT: Baton, Dan L.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Stewart, Timothy A.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zhenli
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3030R1C5
CURRENT APPLICATION NUMBER: US/10/036,342
PRIOR FILING DATE: 2001-12-26
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/112514
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113300
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113430
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/114140
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115552
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/116843
PRIOR FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 60/125774
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125826
PRIOR FILING DATE: 1999-03-24
PRIOR APPLICATION NUMBER: 60/127035
PRIOR FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: 60/127706
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/129122
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/130359
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131272
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/132371
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132379
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132383
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138166
PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: 60/144791
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/146970
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 09/311832
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/380142
```


NUMBER OF SEQ ID NOS: 1198
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 768
LENGTH: 6000
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-860-768

Query Match 6.1%; Score 42.8; DB 10; Length 6000;
Best Local Similarity 50.5%; Pred. No. 0.93;
Matches 104; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 125 TGGACCTGGGGGCTGGGACCTGGGTGCTCCCTGACGAGACGGGGCCCCCTCAGACGACA 184
DB 2687 TGGGCGACGAGATGATGATGCTGGGCTCCGAGATGAGCGCATGGGCGCTGTGATGAGAC 2926
QY 185 GCTGCCCCGGGCTCACCCTGACCGCTCGCATCCCGGACTTTCGCGCGGCGGCGCGG 244
DB 2927 GCATGGGCTCGGTGAGGCGCATGGGCTCCGCGCATTTGAGCGCATGGGCGGCTGG 2986
QY 245 ACCAGATCGACCGCTTTCGATGATGATGAGAGGCGGCGGAGAGGCTGTG 304
DB 2987 ACCAGATGGCTTCAGCATGATGAGGCGAGGCGGCGGAGAGGCGCATGGCTGTGCG 3046
QY 305 GAGTGCACCTGTGCTGTGGGCTTTGGC 330
DB 3047 TGGAGCGCATGGGTGGCGGCGATGGGC 3072

RESULT 13

US-09-747-835A-10
Sequence 10, Application US/09747835A
Patent No. US20020146692A1
GENERAL INFORMATION:
APPLICANT: Yamazaki, Victorla
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Zhou, Ping
APPLICANT: Wang, Dunrui
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Asundi, Vinod
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
FILE REFERENCE: HYS-37CIP
CURRENT APPLICATION NUMBER: US/09/747,835A
PRIOR FILING DATE: 2002-03-08
PRIOR APPLICATION NUMBER: US 09/729,739
PRIOR FILING DATE: 2000-12-04
PRIOR APPLICATION NUMBER: US 09/653,450
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: US 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US 09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: US 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 63
SOFTWARE: Patentin Version 3.0
SEQ ID NO 10
LENGTH: 391
TYPE: DNA
ORGANISM: Homo sapiens
US-09-747-835A-10

Query Match 6.0%; Score 42.2; DB 10; Length 391;
Best Local Similarity 47.7%; Pred. No. 0.56;
Matches 128; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
QY 65 TTCCGGCGCGCTGGCGGAGCTGGCGCTGCCGCGGCTCCCGCCCACTACCACTTCTCTGT 124

DB 109 TGCCCCGCGCTGGCGGCTGCGCTGCGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCTTCC 168
QY 125 TGGACCTGGGGGCTGGGACCTGGGTGCTCCCTGACGAGACGGGGCCCCCTCAGACGACA 184
DB 169 CCGACTGTGAGACCACTGGCTGATGACGACCGGCAAGCATGATGGCTGTGATGACCTTCTCT 228
QY 185 GCTGCCCCGGGCTCACCCTGACCGCTCGCATCCCGGACTTTCGCGCGGCGGCGGCGG 244
DB 229 GCGCGCGGCTGAGCGGCTTCTGATGAGCTTGTGCTGCGGCGGCGGCGGCGGCGGCTTACG 288
QY 245 ACCAGATGACCGCTTTCGATGATGATGAGAGGCGGCGGAGAGGCTGTG 304
DB 289 ACTGTGTAACCTGGCATGATCAAGGAGGTTCGGCCAAAGAGGCGGCGGCTGTGGTGAAG 348
QY 305 GAGTGCACCTGTGCTGTGGGCTTTGGCGGCGAC 335
DB 349 AAGAGTCTGGCGGATGAGATCTACCTCTC 379

RESULT 14

US-09-815-242-7912/C
Sequence 7912, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7912
LENGTH: 1347
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1347)
US-09-815-242-7912

Query Match 6.0%; Score 42.2; DB 10; Length 1347;
Best Local Similarity 46.0%; Pred. No. 0.81;
Matches 145; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 20 CTCTCCAGGAGATGGGGCTGACAGCCGCCCAACTTCTCTGGGTGCTTCCGGCGGCTGG 79
DB 1277 CTCGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGC 1218
QY 80 CGGAGTGGCGGTGCGCGGCGGCTCCCGCCCACTACCAAGTTCTGTTGGACTGGGCGTGC 139

Db 1217 CGGACTGCTTGTCCAGGCTTGCAGCGGCGCACACCGGTTACCGCTTGGCCCAAGTGCCTGG 1158
QY 140 GGCACCTGGTGTCCCTGACGAGCGGGGCCCTTCACAGCGACAGCTGCCCCGGCTCA 199
Db 1157 GCATATCAGCCCGCGCTGACCGCCAGCAAGTACACCGCCCGGAGAGAGCGCGGACGGC 1098
QY 200 CCCTGCACCGCTTGGCGCATCCCGACTTGTGCCCCGGCCCCCGACCAAGATGACCGCT 259
Db 1097 TCGCCAGGCTATGGAAATCGGCGCTTCCGCTCCCGCGGAGAGCAAGCAAGCTTG 1038
QY 260 TCGTCAGATCTGTGACGAGCGCCACGAGCGGAGAGCTGTGGAGTCACTGTGCTC 319
Db 1037 CCGTGAATGTGCGCACACCGCCCTCGATCGCCGCGGAGGCGGCGCGGACGTTGGTGCC 978
QY 320 TGGGCTTGGC 330
Db 977 TTGAATCGTC 967

RESULT 15

US-09-759-143-399
; Sequence 399, Application US/09759143
; Patent No. US20020022248A1

GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Ketter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darlick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 399
; LENGTH: 298
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(298)
; OTHER INFORMATION: n = A,T,C or G
US-09-759-143-399

Query Match

5.9%; Score 42; DB 10; Length 298;

Best Local Similarity 49.8%; Pred. No. 0.57; Mismatches 103; Indels 0; Gaps 0;

QY 125 TGAACCTGGGCGGCGGACCTGTGTCCTGACGAGCGGGGCCCTCAGACAGACA 184
Db 94 TGGGCGACGCGATGATGATGCGTGGCTCCGAGATGAGCGCATGGGCTGTGATGAGC 153
QY 185 GCTGCCCCGGGCTCAACCTGACGCGCTGCGCATCCCGACTTCTGCCCCGGGCCCG 244
Db 154 GCATGGGCTTCGTGGAGGCGCATGGGCTCCGCGCATGAGCGCATGGGCCCGCTGGGCTCG 213
QY 245 ACCAGATGACGCGCTTCGTGCAATCGTGAAGAGCGCAAGCGGAGAGAGGCTGTGG 304
Db 214 ACCACATGGCTCCANCAATTGANGCATGGGCGACCAACCATGGAGCGCATTGGCTTGGCG 273

QY 305 GAGTGCATGTGCTCTGTGGGCTTTGG 329
Db 274 TGGAGCNCATGGGTGCCGCGCATGGC 298

Search completed: November 4, 2002, 10:23:40
Job time : 65 secs

Db 320 GGGACACCTGCGCCGCTGCGCGCGGCGCTGATCCGTCAGGCGCTCCTGCG 379.
Qy 176 ACAGCGACAGTGGCCCGGCGCTGACCTGACACCGGCTGCGATCCCGGACTTCGCGC 235
Db 380 GCGGGAATGGGACACCGGAGACACCGCGTGGAGACCGGATCGCGCGCGCTGCGACCG 439
Qy 236 GCGGCGCGGACGATGACAGCGCTTGTGACAGATCGGAGACGAGCGCAACGCGAGGGAG 295
Db 440 CTCTGCGCAGCCAGCTCCCGGACGCGCTGCGGCTCGCGCACCGCGCGCTCAAGGGCG 499
Qy 296 AGGCTGTGGAGTGCATGCTGCTGCGGCTTGGCGGACATGCGC 339
Db 500 ACGACGTGTGCGCTGCGCGCTGCTGCGGAGCGCGGACGACG 543

RESULT 2

US-08-482-385A-5
Sequence 5, Application US/08482385A
Patent No. 5728561

GENERAL INFORMATION:

APPLICANT: DENOVA, CLAUDIO D.
TITLE OF INVENTION: GENES ENCODING BRACHED CHAIN ALPHA
TITLE OF INVENTION: KETOACID DEHYDROGENASE FROM STREPTOMYCES SPECIES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:

ADDRESSEE: PETER C. RICHARDSON
STREET: 235 EAST 42ND STREET, 20TH FLOOR
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10017-5755

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/482,385A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: SHEYKA, ROBERT F.
REGISTRATION NUMBER: 31,304
REFERENCE/DOCKET NUMBER: PC8346C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-573-1189
TELEFAX: 212-573-1939

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 2728 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-482-385A-5

Query Match 7.8%; Score 55.2; DB 1; Length 2728;
Best Local Similarity 49.6%; Pred. No. 0.0008;

Matches 141; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Qy 56 CTTGGGCTCTTCCGGGCGGCTGCGGAGTGGCGTCCGCGGCTCCCGCCACTAC 115
Db 662 CTTGCGAGTGGCGCGCGCTGCTGAGAGAGCGGAGTGGCTTCCCGACATAC 721
Qy 116 AGTTCCTTTGAGCTGGGCGTGGGACCTGGTCTCCCTGACGAGAGCGGCGCCCTC 175
Db 722 GGGACACCTTGGCGCGCGCTGCGCGGCTGATCCGCTCCGACCTCTGCTGC 781
Qy 176 ACAGCGACAGTGGCGCGGCTGACCTGACACCGGCTGCGGATCCCGGACTTTCGCGC 235
Db 782 GCGGCGACTGGACACCGGAGTACGACCCCGTGGAGACACGCGGCGCGCTGTCACCG 841

Qy 236 GCGGCGCGGACGATGACAGCGCTTGTGACAGATCGTGGACGAGCGCAACGCGAGGGAG 295
Db 842 CTCTGCGACACGAGCTCCCGGACGCGCTGCGGCGCTGCGGACGCGCGCTCAAGGGCG 901
Qy 296 AGGCTGTGGAGTGCATGCTGCTGCGGCTTGGCGGACATGCGC 339
Db 902 ACGACGTGTGCGCTGCGCGCTGCTGCGGCGGAGCGGCGACGACG 945

RESULT 3

US-08-911-853-12
Sequence 12, Application US/08911853
Patent No. 6048710

GENERAL INFORMATION:

APPLICANT: Geritise, Gijbert
APPLICANT: Quax, Wilhelmus J.
TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
TITLE OF INVENTION: EXPRESSION LEVELS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:

ADDRESSEE: Genencor International
STREET: 925 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1013

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,853
FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/699,092
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gjalster, Debra J
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC361-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-846-7620
TELEFAX: 650-845-6504

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:
LENGTH: 642 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-911-853-12

Query Match 6.5%; Score 46.2; DB 3; Length 642;
Best Local Similarity 48.0%; Pred. No. 0.061;

Matches 132; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Qy 4 CCGCGTCTCTGCTGCGCTTCCCGACGAGTGGCGTGGACCGCCCGGACTTCTCTGGG 63
Db 67 CTCGTGTCTCTCTATCATCCTTACGCTGAGCAGACGATGATCTTCTTGGCTG 126
Qy 64 CTTGCGGCGCGGCTGCGGAGTGGCGGCTGCGCGGCTCCCGCCACTTACGATTCCTG 123
Db 127 CTGCGCAGGAGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 186
Qy 124 TTGACCTGGGCGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183
Db 187 TCCATTCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 246
Qy 184 AGTCTGCGGCGGCTGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 243
Db 247 GCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 306
Qy 244 GACGAGATGACCGGCTTGTGCGAGATCGTGAGCA 278

Db 307 AACGCTCCACGGCGATCATCCAGGTCCCGGCGA 341

RESULT 4

US-09-479-409-12
; Sequence 12, Application US/09479409
; Patent No. 6225106
; GENERAL INFORMATION:
; APPLICANT: Gerritse, Gijbert
; APPLICANT: Quax, Wilhelmus J.
; TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
; TITLE OF INVENTION: EXPRESSION LEVELS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1013
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/479,409
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/911,853
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra J
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC361-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-846-7620
; TELEFAX: 650-845-6504
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-479-409-12

Query Match

Best Local Similarity 6.5%; Score 46.2; DB 4; Length 642;
Best Local Similarity 48.0%; Pred. No. 0.061;
Matches 132; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 4 CCCGCTGCTGTCGCTTCCAGCATGGGCGTGGAGCCGCCCACTTCTCCGCTG 63
Db 67 CTGCTGCTGCTGCTATCACCCTGAGCTGAGCAAGAGCATGACTTCTTGGCTG 126
QY 64 CTTCGGGCGGCTGGGCGGAGTGGCTGGCGGCTCCCGCCCACTACCACTTCTG 123
Db 127 CTGCGAGCGAGGCGCGGCGGAGCGCGCCGCGGCGGAGCATGCGGCGGAGCG 186
QY 124 TTGAGCTGGGCTGGGCGGAGTGGCTGGCTGACGAGCGGCGGCGGCGGCGG 183
Db 187 TCCATCCAGGCGCTGAGCATCTGTTGGCAGCGCGCGGCGGCGGCGGAGCG 246
QY 184 AGCTGCCCGGCGCTGACCTGACCGCGCTGGCATGCCGACTTCTGCGGCGGCGG 243
Db 247 GCGGCGGCGGCGGCGGCGGAGCATGACCTGCTGGCGGCGGCGGCGGCGGCG 306
QY 244 GACCAAGTCGACCGCTTCTGTCAGATCGTGAGCA 278
Db 307 AACGCTCCACGGCGATCATCCAGGTCCCGGCGA 341

RESULT 5

US-09-479-453-12
; Sequence 12, Application US/09479453
; Patent No. 6313283

; GENERAL INFORMATION:
; APPLICANT: Gerritse, Gijbert
; APPLICANT: Quax, Wilhelmus J.
; TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
; TITLE OF INVENTION: EXPRESSION LEVELS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1013
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/479,453
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/911,853
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra J
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC361-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-846-7620
; TELEFAX: 650-845-6504
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-479-453-12

Query Match

Best Local Similarity 6.5%; Score 46.2; DB 4; Length 642;
Best Local Similarity 48.0%; Pred. No. 0.061;
Matches 132; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 4 CCCGCTGCTGTCGCTTCCAGCATGGGCGTGGAGCCGCCCACTTCTCCGCTG 63
Db 67 CTGCTGCTGCTGCTATCACCCTGAGCTGAGCAAGAGCATGACTTCTTGGCTG 126
QY 64 CTTCGGGCGGCTGGGCGGAGTGGCTGGCGGCTCCCGCCCACTACCACTTCTG 123
Db 127 CTGCGAGCGAGGCGCGGCGGAGCGCGCCGCGGCGGAGCATGCGGCGGAGCG 186
QY 124 TTGAGCTGGGCTGGGCGGAGTGGCTGGCTGACGAGCGGCGGCGGCGGCGG 183
Db 187 TCCATCCAGGCGCTGAGCATCTGTTGGCAGCGCGCGGCGGCGGCGGAGCG 246
QY 184 AGCTGCCCGGCGCTGACCTGACCGCGCTGGCATGCCGACTTCTGCGGCGGCGG 243
Db 247 GCGGCGGCGGCGGCGGCGGAGCATGACCTGCTGGCGGCGGCGGCGGCGGCG 306
QY 244 GACCAAGTCGACCGCTTCTGTCAGATCGTGAGCA 278
Db 307 AACGCTCCACGGCGATCATCCAGGTCCCGGCGA 341

RESULT 6

US-08-911-853-29/C
; Sequence 29, Application US/08911853
; Patent No. 6048710
; GENERAL INFORMATION:
; APPLICANT: Gerritse, Gijbert

```

; APPLICANT: Quax, Wilhelmus J.
; TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
; TITLE OF INVENTION: EXPRESSION LEVELS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1013
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,853
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,092
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gjalster, Debra J
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC361-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-846-7620
; TELEFAX: 650-845-6504
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17612 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-911-853-29

Query Match          6.5%; Score 46.2; DB 3; Length 17612;
Best Local Similarity 48.0%; Pred. No. 0.14;
Matches 132; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 4 CCCCCTGCTCTGTGCGCCCTTTCACAGATGGGCGGCGAGCCCGCCCACTTCTCTGGGGTG 63
DB 4085 CTCGTGGTGTCTGCTATCATCCTTCAAGCTTCAAGACAGAGCATGATCTTCTGCGCTG 4026
QY 64 CTTCGGGCGGCGCTGCGGAGCTGCGGCTGCCGCGCTCCCGCCCACTTACAGTTCTTG 123
DB 4025 CTGCGACAGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3966
QY 124 TTGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 183
DB 3965 TCCATTCAGCGCGCTGACAGATCTGTTGCGGACGCGCGGCGGCGGCGGCGGCGGCGG 3906
QY 184 AGCTGCGCGGCGCTTACACCTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 243
DB 3905 GCCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3846
QY 244 GACCAAGATGACCGGCTTGTGCGAGATCGTGAGCA 278
DB 3845 AAGCGCTCCAGGCGGATCATCCAGGTGCGCGGCGCA 3811

RESULT 7
US-09-479-409-29/c
; Sequence 29, Application US/09479409
; Patent No. 6225106
; GENERAL INFORMATION:
; APPLICANT: Gerltse, Gijbert
; APPLICANT: Quax, Wilhelmus J.
; TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
; TITLE OF INVENTION: EXPRESSION LEVELS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; COUNTRY: USA
```

```

; ADDRESSEE: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1013
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/479,409
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/911,853
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gjalster, Debra J
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC361-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-846-7620
; TELEFAX: 650-845-6504
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17612 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-479-409-29

Query Match          6.5%; Score 46.2; DB 4; Length 17612;
Best Local Similarity 48.0%; Pred. No. 0.14;
Matches 132; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 4 CCCCCTGCTCTGTGCGCCCTTTCACAGATGGGCGGCGAGCCCGCCCACTTCTCTGGGGTG 63
DB 4085 CTCGTGGTGTCTGCTATCATCCTTCAAGCTTCAAGACAGAGCATGATCTTCTGCGCTG 4026
QY 64 CTTCGGGCGGCGCTGCGGAGCTGCGGCTGCCGCGCTCCCGCCCACTTACAGTTCTTG 123
DB 4025 CTGCGACAGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3966
QY 124 TTGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 183
DB 3965 TCCATTCAGCGCGCTGACAGATCTGTTGCGGACGCGCGGCGGCGGCGGCGGCGGCGG 3906
QY 184 AGCTGCGCGGCGCTTACACCTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 243
DB 3905 GCCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3846
QY 244 GACCAAGATGACCGGCTTGTGCGAGATCGTGAGCA 278
DB 3845 AAGCGCTCCAGGCGGATCATCCAGGTGCGCGGCGCA 3811

RESULT 8
US-09-479-453-29/c
; Sequence 29, Application US/09479453
; Patent No. 6313283
; GENERAL INFORMATION:
; APPLICANT: Gerltse, Gijbert
; APPLICANT: Quax, Wilhelmus J.
; TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
; TITLE OF INVENTION: EXPRESSION LEVELS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
```


QY	19	CGTTCCACACGAGATGGGCGGACAGCCCCCAATTCTCTCGGGGTCTTCGAGGCCGGCTG	78
Db	1215	CCCCCTCGGCCCTCTCTGTCGTCGAGCCCTCTGACCCCTCTGCTCTGCTCTGCCCCCTCTG	1156
QY	79	GCGGGAATGCGGCTGCCCCGGCTCCCCGCCATCAACTCTCTTGGACCTGGGCGTG	138
Db	1155	CCCCCTCTGTGTCCTGAGCCCCCTCTGCCCCCTCTGCTCTCTGCCCCCTCTGCTCTGCCCCCTC	1096
QY	139	CGGACACTGTGTATCCCTGAGACGAGCGCGGGCCCCCTCAACAGGACAGTGGCCCGGGCTC	198
Db	1095	CTGTGCTGCCCCCTCTGCTCTCTGTCGCCCCCTCTGACCCCTCTCTCTGCTCTG	1036
QY	199	ACCGTCACCGCGCTGCGGATCCCGAATTGTGCGCGCGGCCGCCGACCAATGACGCGC	258
Db	1035	CCCCCTCTGTGCTCTGAGCCCCCTCTGCCCCCTCTGACCCCTCTGCTCTGCGCCCTCTCTG	976
QY	259	TTGCTGCAGATCGTG	273
Db	975	CTCTGCCCCCTCTG	961

```

19 CTTTCCACGAGTGGCGTGCAGACCCCAACTTCTCTGGGTCCTTCGCGGCGGCTG 78
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
10
```

D _b	1215	CCCCCTCTGCCCCCTCCCTGCTCTCTGCCCCCTCTGCCCCCTCTCTCTCTCTCTG	1156
Q _Y	79	GCGGAGCTGGCGGTGCGCGGGCTCCCGCCACTACAGATTCTGTGGACTGGGGGT	138
D _b	1155	CCCCCTCTGCTCTGGCCCCCTTCGGCCCCCTGATCTGCTGCCCCCTCTCTCTCTG	1036
Q _Y	139	CGGAGCTGGTCTCTGACGAGAGCCGGGCCCCCTCACAGGAGTGTCCCGGGCTC	198
D _b	1095	CTGCTCTGCGCCCTCTGCTCTCTGCCCCCTGCGCCCTCTCTGCCCCCTCTCTCTG	1036
Q _Y	199	ACCGTCGACCGCTGGCGCATCCCGACTTCTGCCCCCGCGGCCCGGACAGATGACGC	258
D _b	1035	CCCCCTCTGCTCTGCGCCCCCTGCGCCCTCTGCCCCCTGATCTCTCTCTCTG	976
Q _Y	259	TTGCTGAGATGCTG	273
D _b	975	CTCTGCGCCCTCTG	961

```

RESULT 12
US-09-130-114-1
; Sequence 1, Application US/09130114
; Patent No. 5976607
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Dama, Basam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; TITLE OF INVENTION: From Multiple Transfected Epilomes
; FILE REFERENCE: 0867/1D9030S1
; CURRENT APPLICATION NUMBER: US/09/130.114
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: YEBNA
US-09-130-114-1

Query_Match          6.4%; Score 45.4; DB 2; Length 5452;
Best Local Similarity 48.6%; Pred. NO. 0.16;
Matches 124; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY      19  CCTTCCACGAGGATGGGGGTGCAGAGCCGCCCACTCTCTGTGGGTCTTCGGGCGGCTG 78
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       1590  CCCCTCGGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 1649

QY      79  GCGGAGACTGGGCGCTGCGCGGCGCTCCCGCCGACATACAGTACTGTTGGACTGGGCGTG 138
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       1650  CCCCTCTGCTGCTCTGCCCCCTCTGCCCCCTCTGCTGCTGCTGCCCCCTCTGCTG 1709

QY      139  GCGGAGACTGGGTGTCTCCCTACAGGAGGCGGCGCCCTCACAGGAGACTGCCCGGCTC 198
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       1710  CTGCTCTGCCCCCTCTGCTCTCTGCCCCCTCTGCCCCCTCTCTCTCTCTCTCTG 1769

QY      199  ACCCTGCACCGCCTGCGGATCCCGCATCTTGTGGCGCGGCGGCCCGACAGATGACGCG 258
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       1770  CCCCTCGTGTCTCTGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 1829

QY      259  TTGCTGCAATGCTG 273
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       1830  CTCCTGCCCCCTCTG 1844

RESULT 13
US-08-910-647-1/c
; Sequence 1, Application US/08910647
; Patent No. 6251433
; GENERAL INFORMATION:
; APPLICANT: Zuckermann et al.
; TITLE OF INVENTION: Compositions and Methods for

```

```

1  COUNTRY: U.S.A.
2  ZIP: 94608-2916
3
4  COMPUTER READABLE FORM:
5
6  MEDIUM TYPE: Floppy disk
7
8  COMPUTER: IBM PC compatible
9
10 OPERATING SYSTEM: PC-DOS/MS-DOS
11
12 SOFTWARE: PatentIn Release #1.0, Version #1.30
13
14 CURRENT APPLICATION DATA:
15
16 APPLICATION NUMBER: US/09/620,925
17
18 FILING DATE: 21-Jul-2000
19
20 CLASSIFICATION: <Unknown>
21
22 PRIOR APPLICATION DATA:
23
24 APPLICATION NUMBER: 08/910,647
25
26 FILING DATE: <Unknown>
27
28 ATTORNEY/AGENT INFORMATION:
29
30 NAME: Fujita, Sharon M.
31
32 REGISTRATION NUMBER: 38,459
33
34 REFERENCE/DOCKET NUMBER: 1218.002
35
36 TELECOMMUNICATION INFORMATION:
37
38 TELEPHONE: (510) 655-2706
39
40 TELEFAX: (510) 655-3542
41
42 INFORMATION FOR SEQ ID NO: 1:
43
44 SEQUENCE CHARACTERISTICS:
45
46 LENGTH: 9600 base pairs
47
48 TYPE: nucleic acid
49
50 STRANDEDNESS: single
51
52 TOPOLOGY: linear
53
54 MOLECULE TYPE: DNA (genomic)
55
56 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
57
58 US-09-620-925-1
59
60 Query Match 6.4% Score 45.4; DB 4; Length 9600.
61 Best Local Siml|arity 48.6%; Pred NO. 0.18.

```

```

MATCHES 127, CONSERVATIVE 0, MISMATCHES 1317, INDELS 0, GAPS 0
      19 CATTCCACAGCAAGGAGGAGGCGTGCAGGCCCCCAACTTCTCTGAGGTCTTCGCGGCCGCGTG 78
      1261 CCCCCTCGAGCCCTCCTCTCTCTCTGCGCCCTCGCGCCCTCGCTCTCTCTCTGCGCCCTCGTG 1202
      QY 79 GCGGAGCATGCGGCGTGCAGGCGGCGCTCCCGCCACTACAGATTCCTGTTGAACTTGAGCGCTG 138
      DB 1201 CCCCCTCGCTGCCTCGCCCTCTGCGCCCTCTGCTCTGCGCCCTCTCTCTCTGCGCCCTG 1142
      QY 139 CGGACCGCTGATGTCTGCTGACGAGGAGGCGGCGGCCCTCTACAGAGACAGTGGCCCGGCGCTC 198
      DB 1141 CTGCTCTCGTCCCCCTCTCTCTCTCTCTGCGCCCTCTGCGCCCTCTCTCTCTCTCTGCTG 1082
      QY 199 ACCCTGCACGCGCTGCGCGCTCCCGACTTCTGCGCGCGCGGCCCGCCGACAGATCGACCGC 258
      DB 1081 CCCCCTCGCTCTGCGCGCCCTCTGCGCCCTCTGCGCCCTCTGCTCTCTGCGCCCTCTCTG 1022
      QY 259 TTGCTGCAGATCTG 273
      DB 1021 CTCTGCGCCCTCTG 1007

RESULT 15
US-07-884-811-15/c
; Sequence 15, Application US/07884811
; Patent No. 5316921
; GENERAL INFORMATION:
; APPLICANT: Godowsky, Paul J. Lokker, Nathalie A. Mark, Melanie R.
; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 KB floppy disk

```


ALIGNMENTS

(without alignments)
605.685 Million cell updates/sec

Sequence: 1 MGVQPPNFSWVLPGRLAGLA.....GSIETYEQEKAVEQFYQRTK 150

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 10

Post-processing: Listing first 45 summaries

Database : A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1980.DAT.*
 2: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1981.DAT.*
 3: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1982.DAT.*
 4: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1983.DAT.*
 5: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1984.DAT.*
 6: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1985.DAT.*
 7: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1986.DAT.*
 8: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1987.DAT.*
 9: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1988.DAT.*
 10: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1989.DAT.*
 11: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1990.DAT.*
 12: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1991.DAT.*
 13: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1992.DAT.*
 14: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1993.DAT.*
 15: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1994.DAT.*
 16: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1995.DAT.*
 17: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1996.DAT.*
 18: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1997.DAT.*
 19: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1998.DAT.*
 20: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA1999.DAT.*
 21: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA2000.DAT.*
 22: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA2001.DAT.*
 23: /SIDS2/gcgdata/geneseq/geneseq_emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Result No.	Query	Match Length	DB ID	Description
1	SECRET	1	1	

1	150	100.0	150	22	AAB20328	Human protein phos
2	150	100.0	150	22	AAB35275	Human dual specific
3	150	100.0	179	23	ABP43078	Human ovarian anti
4	131	87.3	150	22	AAB73231	Human phosphatase
5	131	87.3	150	22	AA657455	Human acid sequenc
6	131	87.3	150	22	AA657634	Amino acid sequenc
7	131	87.3	150	22	ABBO7846	Human tyrosine phos
8	35	22.3	150	22	AAB73211	Murine phosphatase
9	35	22.3	150	22	AAB35276	Murine dual speci
10	21	14.0	21	22	AAB35276	Human dual specific

XX	AC	AAAB20328 standard; Protein; 150 AA.
XX	DT	AAAB20328;
XX	XX	29-MAY-2001 (first entry)
DE	XX	Human protein phosphatase and kinase protein-7.
KW	KW	Protein phosphatase and kinase protein; PPKP-7; human;
KM	KM	gastrointestinal disorder; immune system disorder;
KX	KX	neurological disorder; cell proliferative disorder; cancer;
OS	OS	diagnosis; therapy.
XX	XX	Homo sapiens.
FH	FH	Key
FT	FT	Location/Qualifiers
FT	FT	Peptide
FT	FT	1..28
FT	FT	/label= Signal_peptide
FT	FT	29..150
FT	FT	/label= Mature_protein
FT	FT	Active-site
FT	FT	93..105
FT	FT	/note= "tyrosine specific protein phosphatase active site signature"
FT	FT	Region
FT	FT	64..107
FT	FT	/note= "Y phosphatase signature"
FT	FT	Domain
FT	FT	93..103
FT	FT	/note= "tyrosine specific protein phosphatase domain signature"
FT	FT	Domain
FT	FT	126..141
FT	FT	/note= "tyrosine specific protein phosphatase domain signature"
FT	FT	Domain
FT	FT	74..125
FT	FT	/note= "tyrosine specific protein phosphatase domain signature"
FT	FT	Modified-site
FT	FT	41
FT	FT	/note= "O-phosphorylated"
FT	FT	Modified-site
FT	FT	43
FT	FT	/note= "O-phosphorylated"
FT	FT	Modified-site
FT	FT	7
FT	FT	/note= "N-glycosylated"
PN	PN	WC200120004-A2.
PD	PD	22-MAR-2001.
XX	XX	14-SEP-2000; 200OWO-US25515.
XX	PR	15-SEP-1999; 99US-0154141.
XX	PA	(INCY-) INCYTE GENOMICS INC.
PI	PI	Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y,
PI	PI	Lu DM;
DR	DR	N-PDSB; AAF30482.
CC	CC	WPI; 2001-244811/25.
PT	PT	The present sequence is that of novel human protein phosphatase and
PT	PT	kinase protein PPKP-7, as predicted from Incyte Clone ID No.
PT	PT	106974CBI (see AAF30482). Tissues that express PPKP-7 (as a
PT	PT	novel human protein phosphatase and kinase proteins for diagnosis,
PT	PT	neurological and cell proliferative disorders -
PT	PT	Claim 1; Page 91; 103pp; English.

fraction of total tissues expressing pPKP-7) include reproductive
CC (0.372), cardiovascular (0.140) and endocrine (0.093). Diseases or
CC conditions associated with tissues expressing pPKP-7 (as a
CC fraction of total tissues expressing pPKP-7) include cancer
CC (0.558), inflammation or trauma (0.233) or cell proliferation
CC (0.209). The encoded protein shows homology to Clon Intestinalis
CC tyrosine phosphatase. The invention provides human pPKP-1 to -11
CC polypeptides (see AAB35275-32) and polynucleotides (see AAF30476-86).
CC It also provides expression vectors, host cells, antibodies, agonists
CC and antagonists, as well as methods for diagnosing, treating or
CC preventing disorders associated with expression of pPKP, including
CC gastrointestinal disorders, immune system disorders, neurological
CC disorders and cell proliferative disorders, including cancer.

SO Sequence 150 AA;
Query Match 100.0%; Score 150; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 1e-145;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVPNPNFSWVLPGRGLALPRLPAHQFLDLGVRHLVSLTERGPPHSDSCPGTLHR 60
DB 1 MGVPNPNFSWVLPGRGLALPRLPAHQFLDLGVRHLVSLTERGPPHSDSCPGTLHR 60
QY 61 LRIPDFCPAPDQIDRFVQIVDEANRGEAVGHCALGFGRTGTMALCYLKERGLAGD 120
DB 61 LRIPDFCPAPDQIDRFVQIVDEANRGEAVGHCALGFGRTGTMALCYLKERGLAGD 120
QY 121 AIAEIRRLRPGSIETYEOKAVFOFYQRTK 150
DB 121 AIAEIRRLRPGSIETYEOKAVFOFYQRTK 150

RESULT 2
AAB35275
ID AAB35275 standard; Protein: 150 AA.
AC AAB35275;
DT 08-MAY-2001 (first entry)
DE Human dual specificity phosphatase DSP-11.
XX
KW Human: dual specificity phosphatase; DSP-11; Duchenne muscular dystrophy;
KW MAP-kinase signalling cascade; cell proliferation; cancer; allergy;
KW graft-versus-host disease; autoimmune disease; metabolic disease.
XX
OS Homo sapiens.
XX
PN MO200105983-A1.
PD 25-JAN-2001.
PF 19-JUL-2000; 2000MO-US19710.
PR 20-JUL-1999; 990S-0144557.
PA (CEPT-) CEPTYR INC.
PI Luche RM, Wei B;
XX
DR WPI: 2001-147348/15.
DR N-PSDB; AAF27958.
XX
PT Novel dual specificity phosphatase-11 dephosphorylating activated
PT mitogen-activated protein kinase, is used to identify agents inhibiting
PT the enzyme activity and modulate cell proliferation -
XX
PS Claim 1; Fig 2; 65pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC human dual-specificity phosphatase DSP-11. The protein is involved in the
CC MAP-kinase signalling cascade. The sequences can be used in the diagnosis

CC and treatment of Duchenne muscular dystrophy, cancer, graft-versus-host
CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal
CC cell growth and abnormal cell proliferation. The present sequence is the
CC human DSP-11 protein.

SO Sequence 150 AA;
Query Match 100.0%; Score 150; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 1e-145;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVPNPNFSWVLPGRGLALPRLPAHQFLDLGVRHLVSLTERGPPHSDSCPGTLHR 60
DB 1 MGVPNPNFSWVLPGRGLALPRLPAHQFLDLGVRHLVSLTERGPPHSDSCPGTLHR 60
QY 61 LRIPDFCPAPDQIDRFVQIVDEANRGEAVGHCALGFGRTGTMALCYLKERGLAGD 120
DB 61 LRIPDFCPAPDQIDRFVQIVDEANRGEAVGHCALGFGRTGTMALCYLKERGLAGD 120
QY 121 AIAEIRRLRPGSIETYEOKAVFOFYQRTK 150
DB 121 AIAEIRRLRPGSIETYEOKAVFOFYQRTK 150

RESULT 3
ABP43078
ID ABP43078 standard; Protein: 179 AA.
AC ABP43078;
DT 22-AUG-2002 (first entry)
DE Human ovarian antigen HVCAJ27, SEQ ID NO:4210.
XX
KW Human: ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive.
XX
OS Homo sapiens.
XX
PN MO200200677-A1.
PD 03-JAN-2002.
PF 07-JUN-2001; 2001MO-US18569.
PR 07-JUN-2000; 2000UG-209467P.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Birse CE, Rosen CA;
XX
DR WPI: 2002-147878/19.
DR N-PSDB; ABQ56155.
XX
PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -
XX
PS Claim 11; SEQ ID NO 4210; 2922pp; English.
XX
CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;
 PI Senoo C, Nezu J;
 DR WPI; 2001-564736/63.
 XX
 PT New genes encoding protein kinase and protein phosphatase, useful for
 PT identifying modulators which can be used to treat human or animal
 PT disorders associated with the expression or function of these enzymes -
 XX
 PS Example 4; Page 297-298; 336pp; Japanese.
 XX
 CC The specification describes human protein kinase/protein phosphatases.
 CC The polypeptides are expected to participate in signal transduction
 CC in cells. The kinase phosphatases are connected with intracellular
 CC signalling pathways. Antisense oligonucleotides and compounds
 CC identified by screening (agonists or antagonists) can be used to
 CC treat human or animal disorders associated with the expression
 CC or function of the protein. In addition, the polypeptides may be used
 CC as target molecules for drug development. The present sequence
 CC represents a polypeptide, used in the course of the invention.
 XX
 SQ Sequence 150 AA;
 XX
 Query Match 87.3%; Score 131; DB 22; Length 150;
 Best Local Similarity 100.0%; Pred. No. 3.2e-126;
 Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGVPNPNFVWLPGRLAGLALPRPAHYQFLDLGVRHLVSLTERGPPHSDSCPELTILHR 60
 DB 1 MGVPNPNFVWLPGRLAGLALPRPAHYQFLDLGVRHLVSLTERGPPHSDSCPELTILHR 60
 QY 61 LRIIDFCPPADQIDRFVQIYDEANRGEAVGHCALFGRTGIMLACYLKERGLAAGD 120
 DB 61 LRIIDFCPPADQIDRFVQIYDEANRGEAVGHCALFGRTGIMLACYLKERGLAAGD 120
 QY 121 AIAEIRRLRPG 131
 DB 121 AIAEIRRLRPG 131
 XX
 RESULT 6
 AAG67634
 ID AAG67634 standard; Protein: 150 AA.
 XX
 AC AAG67634;
 XX
 DT 26-NOV-2001 (first entry)
 XX
 DE Amino acid sequence of a human protein.
 XX
 KW Human; protein kinase; protein phosphatase; signal transduction.
 XX
 OS Homo sapiens.
 XX
 PN WO200109316-A1.
 XX
 PD 08-FEB-2001.
 XX
 PF 28-JUL-2000; 2000WO-JP05061.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 18-OCT-1999; 99US-0159590.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 17-FEB-2000; 2000US-0183322.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;
 PI Senoo C, Nezu J;

XX
 DR WPI; 2001-570286/64.
 XX
 PT New genes encoding proteins with protein kinase/protein phosphatase
 PT activity, useful in the diagnosis and treatment of diseases -
 XX
 PS Example 4; Page 197-198; 233pp; Japanese.
 XX
 CC The specification describes human protein kinase/protein phosphatases.
 CC It is expected that the protein kinase/protein phosphatase gene
 CC participates in signal transduction in cells. The protein
 CC kinase/protein phosphatase polypeptides and polynucleotides are
 CC useful for developing diagnostics and treatment agents for human
 CC and animal diseases. The protein kinase/protein phosphatase polypeptides
 CC are useful as target molecules in designing novel drugs. The protein
 CC kinase/protein phosphatase polynucleotides are useful as a source of
 CC probes and primers, which may be used to isolate homologous sequences.
 CC The present sequence represents a human protein, which is used in the
 CC course of the invention.
 XX
 SQ Sequence 150 AA;
 XX
 Query Match 87.3%; Score 131; DB 22; Length 150;
 Best Local Similarity 100.0%; Pred. No. 3.2e-126;
 Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGVPNPNFVWLPGRLAGLALPRPAHYQFLDLGVRHLVSLTERGPPHSDSCPELTILHR 60
 DB 1 MGVPNPNFVWLPGRLAGLALPRPAHYQFLDLGVRHLVSLTERGPPHSDSCPELTILHR 60
 QY 61 LRIIDFCPPADQIDRFVQIYDEANRGEAVGHCALFGRTGIMLACYLKERGLAAGD 120
 DB 61 LRIIDFCPPADQIDRFVQIYDEANRGEAVGHCALFGRTGIMLACYLKERGLAAGD 120
 QY 121 AIAEIRRLRPG 131
 DB 121 AIAEIRRLRPG 131
 XX
 RESULT 7
 ABB07846
 ID ABB07846 standard; Protein: 150 AA.
 XX
 AC ABB07846;
 XX
 DT 03-JUL-2002 (first entry)
 XX
 DE Human tyrosine phosphatase-like enzyme.
 XX
 KW Tyrosine phosphatase-like enzyme; antidiabetic; anorectic; cytostatic;
 KW cardiant; metabolic; immunomodulator; antiparkinsonian; anticonvulsant;
 KW cerebroprotective; neuroprotective; nootropic; neuroleptic; anti-HIV;
 KW antiarrhythmic; hypotensive; vulnerary; gene therapy; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200220747-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 05-SEP-2001; 2001WO-EP10205.
 XX
 PR 11-SEP-2000; 2000US-231568P.
 PR 06-DEC-2000; 2000US-251403P.
 XX
 PA (FARB) BAYER AG.
 XX
 Key Location/Qualifiers
 FT 92..108
 FT Region /note- "the encoding nucleotide fragment for the above
 FT residues is not indicated in the corresponding
 FT DNA sequence"

PI Kossida S;
XX
DR WPI; 2002-339803/37.
DR N-PSDB; ABL40806.
XX
PT New human tyrosine phosphatase-like enzyme polypeptide, regulators of
PT which are useful for preventing, treating diabetes, obesity, cancer,
PT cardiovascular and pulmonary diseases
XX
PS Claim 18; Fig 2; 117pp; English.
XX
CC The invention relates to a purified human tyrosine phosphatase-like
CC enzyme polypeptide. The enzyme can be expressed by standard recombinant
CC methodology. The tyrosine phosphatase-like enzyme and encoding
CC polynucleotides are useful for screening for modulators which are used
CC for treating a tyrosine phosphatase-like enzyme dysfunction related
CC disease such as a central nervous system (CNS) disorder, diabetes,
CC obesity, chronic obstructive pulmonary disease, cardiovascular disease,
CC cancer, anorexia, cachexia, wasting disorders, appetite suppression, and
CC other eating disorder such as bulimia. CNS disorders include brain
CC injuries, Parkinson's disease, dementia, multiple sclerosis, stroke,
CC Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease,
CC Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, pain associated with CNS disorders
CC and cardiovascular diseases include myocardial infarction, ischaemic
CC diseases of the heart, atrial and ventricular arrhythmia, hypertensive
CC vascular diseases and peripheral vascular diseases. The enzyme is useful
CC in diagnostic assays for detecting diseases and abnormalities or
CC susceptibility to diseases or abnormalities related to the presence of
CC mutations in the encoding nucleic acid sequences. The present sequence
CC represents the human tyrosine phosphatase-like enzyme polypeptide.
XX
SQ Sequence 150 AA;
XX
Query Match 87.3%; Score 131; DB 23; Length 150;
Best Local Similarity 100.0%; Pred. No. 3,2e-126;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MGVPNPFNSVVLPGRLAGLALPRLPAHYQFLDLGVNHLVSLTERGPHSDSCGLTLNR 60
DB 1 MGVPNPFNSVVLPGRLAGLALPRLPAHYQFLDLGVNHLVSLTERGPHSDSCGLTLNR 60
XX
QY 61 LRIPLDFCPAPDQIDRFVQIVDEANARGEAVGVHICALGFRTGTMACYLVKERGLAAGD 120
DB 61 LRIPLDFCPAPDQIDRFVQIVDEANARGEAVGVHICALGFRTGTMACYLVKERGLAAGD 120
XX
QY 121 AIAEIRRLRPG 131
DB 121 AIAEIRRLRPG 131
XX
RESULT 8
AAB73211
ID AAB73211 standard; Protein; 150 AA.
XX
AC AAB73211;
XX
DT 11-MAY-2001 (first entry)
XX
DE Murine phosphatase AA023073_m.
XX
KW phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
KW congenital muscle disorder; Papillon-Leleuvre syndrome; Cowden disease;
KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
KW schizophrenia; hamartoma.
XX
OS Mus sp.
XX
PN WO200112819-A2.
XX
PD 22-FEB-2001.
XX

PF 11-AUG-2000; 2000MO-DS22158.
XX
PR 13-AUG-1999; 99US-0149005.
XX
PA (SUG-) SUGEN INC.
XX
PI Florman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;
XX
DR WPI; 2001-211226/21.
DR N-PSDB; AAF63563.
XX
PT New protein phosphatase polypeptide for diagnosing and treating
PT phosphatase related disorders such as cancer, schizophrenia, cardiac
PT dysfunction and/or vascular disorders
XX
PS Claim 6; Fig 5; 138pp; English.
XX
CC The present invention relates to phosphatase proteins and coding
CC sequences. The present sequence is one such phosphatase. Phosphatases are
CC enzymes that catalyse the dephosphorylation of proteins modified by
CC phosphorylation of serine, threonine or tyrosine residues. The
CC phosphatases are useful for treating a variety of diseases: for example
CC cancer e.g. breast, uterine, prostate, head, neck, lung cancers,
CC synovial sarcomas, renal cell carcinoma, non-small cell lung cancer,
CC hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,
CC glioblastoma, colorectal cancer and thyroid cancer, pathophysiological
CC hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,
CC congenital muscle disorders, Papillon-Leleuvre syndrome, Cowden disease,
CC ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan
CC Zonana syndrome, schizophrenia and hamartomas.
XX
SQ Sequence 150 AA;
XX
Query Match 23.3%; Score 35; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 9,6e-28;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 80 IVDENARGEAVGVHICALGFRTGTMACYLVKER 114
DB 80 IVDENARGEAVGVHICALGFRTGTMACYLVKER 114
XX
RESULT 9
AAB35276
ID AAB35276 standard; Protein; 150 AA.
XX
AC AAB35276;
XX
DT 08-MAY-2001 (first entry)
XX
DE Murine dual specificity phosphatase DSP-11.
XX
KW Mouse; dual specificity phosphatase; DSP-11; Duchenne muscular dystrophy;
KW MAP-kinase signalling cascade; cell proliferation; cancer; allergy;
KW graft-versus-host disease; autoimmune disease; metabolic disease.
XX
OS Mus sp.
XX
PN WO200105983-A1.
XX
PD 25-JAN-2001.
XX
PF 19-JUL-2000; 2000MO-US19710.
XX
PR 20-JUL-1999; 99US-0144557.
XX
PA (CEPT-) CEPTYR INC.
XX
PI Lucche RM, Wei B;
XX
DR WPI; 2001-147348/15.
DR N-PSDB; AAF27959.
XX

PT Novel dual specificity phosphatase-11 dephosphorylating activated
PT mitogen-activated protein kinase, is used to identify agents inhibiting
PT the enzyme activity and modulate cell proliferation -
XX
XX Example 3; Fig 6; 65pp; English.
PS
CC The present invention provides the protein and coding sequences of the
CC human dual-specificity phosphatase DSP-11. The protein is involved in the
CC MAP-kinase signaling cascade. The sequences can be used in the diagnosis
CC and treatment of Duchenne muscular dystrophy, cancer, graft-versus-host
CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal
CC cell growth and abnormal cell proliferation. The present sequence is the
CC mouse DSP-11 protein.
XX
SQ Sequence 150 AA;

Query Match 23.3%; Score 35; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 9.6e-28;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 IYDEANRGEAVGVHCAIGFGRTGTMACYLYKER 114
|||
DB 80 IYDEANRGEAVGVHCAIGFGRTGTMACYLYKER 114

RESULT: 10
AAB35273
ID AAB35273 standard; peptide: 21 AA.
XX
AC AAB35273;
XX
DT 08-MAY-2001 (first entry)
XX
DE Human dual specificity phosphatase DSP-11 active site domain.
XX
KM Human: dual specificity phosphatase; DSP-11; Duchenne muscular dystrophy;
KM MAP-kinase signaling cascade; cell proliferation; cancer; allergy;
KM graft-versus-host disease; autoimmune disease; metabolic disease.
XX
XX Homo sapiens.
OS
XX WO200105983-A1.
PN
XX 25-JAN-2001.
PD
XX 19-JUL-2000; 2000WO-US19710.
PF
XX 20-JUL-1999; 99US-0144557.
PR
XX (CEPT-) CEPTYR INC.
PA
XX Luche RM, Wei B;
PI
XX WPI; 2001-147348/15.
DR
XX
XX Novel dual specificity phosphatase-11 dephosphorylating activated
PT mitogen-activated protein kinase, is used to identify agents inhibiting
PT the enzyme activity and modulate cell proliferation -
XX
XX Disclosure; Page 8; 65pp; English.
PS
XX The present invention provides the protein and coding sequences of the
CC human dual-specificity phosphatase DSP-11. The protein is involved in the
CC MAP-kinase signaling cascade. The sequences can be used in the diagnosis
CC and treatment of Duchenne muscular dystrophy, cancer, graft-versus-host
CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal
CC cell growth and abnormal cell proliferation. The present sequence is the
CC human DSP-11 active site domain.
XX
SQ Sequence 21 AA;

Query Match 14.0%; Score 21; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.7e-14;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 91 VGVHCAIGFGRTGTMACYLY 111
|||
DB 1 VGVHCAIGFGRTGTMACYLY 21

Search completed: November 4, 2002, 12:57:59
Job time : 33 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 2002, 07:55:42 ; Search time 2068 Seconds

(without alignments)
5536.852 Million cell updates/sec

Title: US-09-619-380-1

Perfect score: 707
Sequence: 1 tgaaccgcgtctcgtgcc.....aaaaaaaaaaaaaaaaa 707

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estrov:*
6: em_estropl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	694.2	98.2	808	12	BG823348
2	690.8	97.7	709	14	BM676758
3	683.8	96.7	740	13	BT761974
4	676.8	95.7	747	14	BO679163
5	675.8	95.7	747	14	BO684685
6	675.4	95.3	1127	14	BM924391

7	671.6	95.0	709	10	BE263806	BE263806	601194169
8	645.8	91.3	725	13	BT598433	BT598433	603246943
9	621.4	87.9	904	14	BO892874	BO892874	AGENCOURT
10	608.6	86.1	712	12	BF970174	BF970174	602273385
11	592.6	83.8	828	12	BG831257	BG831257	602766474
12	587.4	83.1	639	12	BG823185	BG823185	602726543
13	583.8	82.6	602	9	AI681916	AI681916	tx50f02.x
14	555.6	78.6	931	14	BQ214823	BQ214823	AGENCOURT
15	551	77.9	584	13	BT761318	BT761318	603044078
16	542.8	76.8	634	13	BT116133	BT116133	602866343
17	508.8	72.0	1059	14	BO680732	BO680732	AGENCOURT
18	508.6	71.9	553	10	BE264444	BE264444	601191757
19	500.2	70.7	1066	14	BO682193	BO682193	AGENCOURT
20	495.4	70.1	1101	13	BT165135	BT165135	603051358
21	489	69.2	604	12	BT075013	BT075013	602686517
22	487.8	69.0	1008	14	BO892811	BO892811	AGENCOURT
23	486.4	68.8	680	13	BT165377	BT165377	603050403
24	486.2	68.5	493	10	BE220956	BE220956	hu02b01.x
25	484	68.5	485	9	AI1739584	AI1739584	w135404.x
26	477.8	67.6	481	9	AI302179	AI302179	GR58C06.x
27	476.4	67.4	500	14	BM853970	BM853970	K-EST0136
28	475.4	67.2	477	9	AI597715	AI597715	tu51g08.x
29	472.2	66.8	477	9	AI689729	AI689729	tx30f02.x
30	462	65.3	1255	14	BO678493	BO678493	AGENCOURT
31	454.8	64.3	650	12	BE728244	BE728244	601563392
32	453.8	64.2	457	12	BF062579	BF062579	7b61h10.x
33	451.6	63.9	458	9	AI432132	AI432132	tg77a08.x
34	446	63.1	448	9	AI363485	AI363485	qy69f11.x
35	441.8	62.5	445	9	AI309010	AI309010	q90b007.x
36	434.8	61.5	441	10	BE671361	BE671361	7e49h11.x
37	432	61.1	440	10	AM664958	AM664958	h198d06.x
38	429	60.7	453	9	AI616223	AI616223	au45g10.y
39	428.2	60.6	521	13	BT847550	BT847550	469416 MA
40	414	58.6	450	14	BM179897	BM179897	UI-E-EJ0-
41	412	58.3	412	9	AA479435	AA479435	zv14e10.r
42	411.4	58.2	475	9	AI341528	AI341528	qx92d02.x
43	408.6	57.8	417	9	AA101788	AA101788	2x96c10.s
44	406.2	57.5	754	14	BO181810	BO181810	UI-H-EJ0-
45	402.8	57.0	406	9	AI223430	AI223430	q954b02.x

ALIGNMENTS

RESULT 1
LOCUS BG823348 808 bp mRNA linear EST 22-MAY-2001
DEFINITION 602726743F1 NIH_MGC_15 Homo sapiens CDNA clone IMAGE:4866058 5',
ACCESSION BG823348
VERSION BG823348.1 GI:14170935
KEYWORDS EST.
SOURCE Human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 808)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L10CM1732 row: f column: 11
High quality sequence stop: 745.
Location/Qualifiers
1. 808

FEATURES
source

```

/organism="Homo sapiens"
/dx_xref="taxon:9606"
/clone="IMAGE:4866058"
/clone_id="N1H.MGC-15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pOTB1; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT      172 a      256 c      242 g      138 t
ORIGIN
Query Match      98.2%; Score 694.2; DB 12; Length 808;
Best Local Similarity 98.9%; Pred. No. 6.3e-99;
Matches 699; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 TGACCCGCTGCTCTGTCGCTTTCCTCCAGATGGGCGTGCAGACCCCACTTCCTG 60
DB 51 TGACCCGCTGCTCTGTCGCTTTCCTCCAGATGGGCGTGCAGACCCCACTTCCTG 110
QY 61 GTGCTTCGGGCGGCTGGCGGAGTGGCGCTGCGCGGCTCCCGCCACTACAGTTC 120
DB 111 GTGCTTCGGGCGGCTGGCGGAGTGGCGCTGCGCGGCTCCCGCCACTACAGTTC 170
QY 121 CTGTTGACCTGGGCGTGGCGGACCTGCTGCTCCGACGAGGCGGCGGCGGCTCAGC 180
DB 171 CTGTTGACCTGGGCGTGGCGGACCTGCTGCTCCGACGAGGCGGCGGCGGCTCAGC 230
QY 181 GACAGCTGCGGCGGCTGACCTGACCTGCGGCGATCCCGACTTCTGCGCGGCGGCG 240
DB 231 GACAGCTGCGGCGGCTGACCTGACCTGCGGCGATCCCGACTTCTGCGCGGCGGCG 290
QY 241 CCCGACGATGACGCGCTTGTGCAATGCTGACGAGGCGCAACGCGAGGAGAGCT 300
DB 291 CCCGACGATGACGCGCTTGTGCAATGCTGACGAGGCGCAACGCGAGGAGAGCT 350
QY 301 GTGGGAGTGCACCTGCTGCTGGGCTTGGCGGACCTGCGGCGATGCTGCTGCTACTG 360
DB 351 GTGGGAGTGCACCTGCTGCTGGGCTTGGCGGACCTGCGGCGATGCTGCTGCTACTG 410
QY 361 GTGAAGAGCGGCGGCTTGGCTGAGAGATGCCATGCTGAATCCGACGACTACGACC 420
DB 411 GTGAAGAGCGGCGGCTTGGCTGAGAGATGCCATGCTGAATCCGACGACTACGACC 470
QY 421 GGTCCATCGAGACCTATGACGAGAGAGAGACAGTCTTCCAGTCTACGAGAGAGAA 480
DB 471 AGCTCCATCGAGACCTATGACGAGAGAGAGAGACAGTCTTCCAGTCTACGAGAGAA 530
QY 481 TAAAGGCGCTTAACTTACCTTACAGAGGCTCACTCCCTCCGCAATGTTGTGATGG 540
DB 531 TAAAGGCGCTTAACTTACCTTACAGAGGCTCACTCCCTCCGCAATGTTGTGATGG 590
QY 541 GCCAGAGATGAAGGAGAGTGAATTAAGTATTAACCTTACGCTTCCATTTGCTGAGA 600
DB 591 GCCAGAGATGAAGGAGAGTGAATTAAGTATTAACCTTACGCTTCCATTTGCTGAGA 650
QY 601 CACTGAAGTACCCCAACCTTACAGGAGGCTGCTGATGAGAGGAGGCTTACTGCTTT 660
DB 651 CACTGAAGTACCCCAACCTTACAGGAGGCTGCTGATGAGAGGAGGCTTACTGCTTT 710
QY 661 GTTGAATAATGATTTTACGACCAAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTT 707
DB 711 GTTGAATAATGATTTTACGACCAAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTT 757

```

```

RESULT 2
BM676758/c
LOCUS      BM676758      709 bp      mRNA      linear      EST 27-FEB-2002

```

```

DEFINITION      UI-E-EJ0-ahu-1-09-0-UI.s2 UI-E-EJ0 Homo sapiens cDNA clone
ACCESSION      UI-E-EJ0-ahu-1-09-0-UI 3', mRNA sequence.
VERSION      BM676758
KEYWORDS      BM676758.1 GI:18986654
SOURCE      EST.
ORGANISM      human.
REFERENCE      Homo sapiens
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL      1 (bases 1 to 709)
MEDLINE      Bonaldo,M.F., Lennon,G. and Soares,M.B.
COMMENT      Normalization and subtraction: two approaches to facilitate gene
              discovery
              Genome Res. 6 (9), 791-806 (1996)
              9704447
              Contact: Soares, MB
              Program for Rat Gene Discovery and Mapping
              University of Iowa
              451 Eckstein Medical Research Building Iowa City, IA 52242, USA
              Tel: 319 335 8250
              Fax: 319 335 9565
              Email: mssoares@blue.weeg.uiowa.edu
              Tissue Procurement: Dr. Gregg Hageman
              cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: Researchers may obtain clones from Research
              Genetics (www.resgen.com).
              Sex: Male
              Polymorphism: M13 Forward
              Location/Qualifiers
              1. 709
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="UI-E-EJ0-ahu-1-09-0-UI"
              /clone_id="UI-E-EJ0"
              /tissue_type="fetal eyes, lens, eye anterior segment,
              optic nerve, retina, Retina Foveal and Macular, RPE and
              Choroid"
              /dev_stage="fetal and adult"
              /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
              /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
              modified polylinker; Site:1: EcoR I; Site:2: Not I;
              UI-E-EJ0 is a subtracted cDNA library constructed
              according to Bonaldo, Lennon and Soares, Genome Research,
              6:791-806, 1996. First strand cDNA synthesis was primed
              with an oligo-dT primer containing a Not I site. Double
              stranded cDNA was ligated to an EcoR I adaptor, digested
              with Not I, and cloned directionally into pT73-Pac
              vector. The oligonucleotide used to prime the synthesis of
              first-strand cDNA contains a library tag sequence that is
              located between the Not I site and the (dfr)18 tail. The
              sequence tags for this library are: fetal eyes, AGATATACAA
              ; lens, CCAATAGCA; eye anterior segment, AATGCCGAT;
              optic nerve, CCATTAAGTC; retina, CCGCG; Retina Foveal and
              Macular, GTCC; RPE and Choroid, ACCCTA. This library was
              created for the program, Gene Discovery in the Visual
              System, supported by National Eye Institute (NEI).
              TAG-LIB-UI-E-EJ0
              TAG-TISSUE-Foveal and Macular Retina
              TAG-SEQ-GTCC
BASE COUNT      135 a      203 c      227 g      144 t
ORIGIN
Query Match      97.7%; Score 690.8; DB 14; Length 709;
Best Local Similarity 99.6%; Pred. No. 2.3e-98;
Matches 703; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 TGACCCGCTGCTCTGTCGCTTTCCTCCAGATGGGCGTGCAGACCCCACTTCCTG 60
DB 705 TGACCCGCTGCTCTGTCGCTTTCCTCCAGATGGGCGTGCAGACCCCACTTCCTG 647
QY 61 GTGCTTCGGGCGGCTGGCGGAGTGGCGCTGCGCGGCTCCCGCCACTACAGTTC 120

```

Db	646	GTCCTTCCGGGCGGCTGGCGAGACTGGGCGTCCGCCGCTCCCCACCACTACAGTTT	587
Oy	121	CTGTGGACCTGGGGCGCTGGGGACCTGNGTGTCCTCCTACAGAGCGGGGCCCCCTCACAGC	180
Db	586	CTGTGGACCTGGGGCGCTGGGGACCTGGTGTCCTCCTACAGAGCGGGGCCCCCTCACAGC	527
Oy	181	GACAGCTGCCCGGCTTCACCCCTGCACCCTTCGCGCATCCCAGACTTTCGCCGCGGGCC	240
Db	526	GACAGCTGCCCGGCTTCACCCCTGCACCCTTCGCGCATCCCAGACTTTCGCCGCGGGCC	467
Oy	241	CCCGACCAATGCAGCCGCTTCGTGCAGATCTGGACGAGGCCAACGACGGGGAGAGCT	300
Db	466	CCCGACCAATGCAGCCGCTTCGTGCAGATCTGGACGAGGCCAACGACGGGGAGAGCT	407
Oy	301	GTGGGAGTCACTGTCGTGGGCTTTGGCCGCACTGGAGGCATGCTGGCCCTGTACCTG	360
Db	406	GTGGGAGTCACTGTCGTGGGCTTTGGCCGCACTGGAGGCATGCTGGCCCTGTACCTG	347
Oy	361	GTGAAGGAGCGGGGCTTTGGCTCAGAGAGATGCCATTGCTGAATCCGACGACTACGACC	420
Db	346	GTGAAGGAGCGGGGCTTTGGCTCAGAGAGATGCCATTGCTGAATCCGACGACTACGACC	287
Oy	421	GGCTCCATCGAACAACCTATATAGCAGAGAAAGACATGTTCTCACTTACAGCAAGAAA	480
Db	286	GGCTCCATCGAACAACCTATATAGCAGAGAAAGACATGTTCTCACTTACAGCAAGAAA	227
Oy	481	TAAAGGGCCCTTAGTACCTCTTACACAGGCGCTCACACCCCTTCCCATTTGNCATGGG	540
Db	226	TAAAGGGCCCTTAGTACCTCTTACACAGGCGCTCACACCCCTTCCCATTTGNCATGGG	167
Oy	541	GCCAAGATGAAAGGAAAGTGAATAAGTATTAACCCCTTAGCTCCCATTTGGCTGAAGA	600
Db	166	GCCAAGATGAAAGGAAAGTGAATAAGTATTAACCCCTTAGCTCCCATTTGGCTGAAGA	107
Oy	601	CACCTAAGTAGCCACCCCTTCGACGAGGCTCTGATTTGAAGGGGAGGCTGTACGCTTT	660
Db	106	CACCTAAGTAGCCACCCCTTCGACGAGGCTCTGATTTGAAGGGGAGGCTGTACGCTTT	47
Oy	661	GTTGAATAAATGATTTTACGAAACCAAAAAAAAAAAAAA 706	
Db	46	GTTGAATAAATGATTTTACGAAACCAAGAAAAAAAAAAAAA 1	
RESULT 3			
LOCUS	Bt761974	740 bp mRNA linear EST 25-SEP-2001	
DEFINITION		603046920P1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5189045 5',	
ACCESSION	Bt761974	mRNA sequence.	
VERSION	Bt761974.1	GI:15735552	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	DNA Sequencing by: Inyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.llnl.gov		
	Plate: LRAM1472 row: h column: 06		
	High quality sequence stop: 720.		
FEATURES			
source	Location/Qualifiers		
	I..740		

Query Match	Similarity	Score	DB	Length
Matches 696; Conservative	99.6%;	Pred. No. 2.7e-97;	Indels 1;	Gaps 1;
BASE COUNT	139 a	239 c	222 g	140 t
ORIGIN				
<p> "/organism="Homo sapiens" "/db_xref="taxon:9606" "/clone="IMAGE:5189045" "/clone_11b="NIH_MGC_116" "/lab_host="DH10B" "/note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is 0.19g-df primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH_MGC Library." </p>				
Query 1	TSACCCGCTGTCCTCTGTCCTTCCACAGAGAGGGCGTACAGCCCCCAACTTCCTCGG	60		
Db 41	TSACCCGCTGTCCTCTGTCCTTCCACAGAGAGGGCGTACAGCCCCCAACTTCCTCGG	100		
QY 61	GTGCTTCGGGGCGGGCTGGCGGAGTGGCGCTGCGCGGCTCCCGCCCACTACAGTTC	120		
Db 101	GTGCTTCGGGGCGGGCTGGCGGAGTGGCGGCTGCGCGGCTCCCGCCCACTACAGTTC	160		
QY 121	CTGTTGACCTGGGGGTGGCGGACCTGGGTGTCCTTGACGAGACGGCGGGCCCTCAGAC	180		
Db 161	CTGTTGACCTGGGGGTGGCGGACCTGGGTGTCCTTGACGAGACGGCGGGCCCTCAGAC	220		
QY 181	GACAGCTGCCCCGGGCTTACCTGCACGCGCTGCGCATCCCGACTTCGCGCGCGGC	240		
Db 221	GACAGCTGCCCCGGGCTTACCTGCACGCGCTGCGCATCCCGACTTCGCGCGCGGC	280		
QY 241	CCCGACCAAGTGGACCGGCTTGGAGATCGTGGACGAGGCGCAAGCGAGGGGAGAGGT	300		
Db 281	CCCGACCAAGTGGACCGGCTTGGAGATCGTGGACGAGGCGCAAGCGAGGGGAGAGGT	340		
QY 301	GTGGAGTGACACTGTGCTGTGGGCTTTGGCCGACATGGACACCATGCTGCTTACCTG	360		
Db 341	GTGGAGTGACACTGTGCTGTGGGCTTTGGCCGACATGGACACCATGCTGCTTACCTG	400		
QY 361	GTGAAGAGAGCGGGGCTTGGCTGACAGAGATGCCATTGCTGAATCCGACGACTACGACC	420		
Db 401	GTGAAGAGAGCGGGGCTTGGCTGACAGAGATGCCATTGCTGAATCCGACGACTACGACC	460		
QY 421	GGCCTCATCGAGACCTATAGAGAGAGAAAGACAGTCTTCAGATTTCACAGGAGAGAA	480		
Db 461	GGCCTCATCGAGACCTATAGAGAGAGAAAGACAGTCTTCAGATTTCACAGGAGAGAA	520		
QY 481	TAAAGGGCCTTAGTACCTTACAGAGGCGCTCACTCCCTTCCCATGTTGTCATGGG	540		
Db 521	TAAAGGGCCTTAGTACCTTACAGAGGCGCTCACTCCCTTCCCATGTTGTCATGGG	580		
QY 541	GCCAGAGTGAAGGAGACCTGAAGCTAAAGTATTAACCCCTTAGGTCCTCATTGGCGAAG	600		
Db 581	GCCAGAGTGAAGGAGACCTGAAGCTAAAGTATTAACCCCTTAGGTCCTCATTGGCGAAG	640		
QY 601	CACGTAGATAGCCACCCCTGCAGGACAGTCTCGATTGAAGGGAGGCGTTGATC-TGCTT	659		
Db 641	CACGTAGATAGCCACCCCTGCAGGACAGTCTCGATTGAAGGGAGGCGTTGATC-TGCTT	700		
QY 660	TGTTGAATTAATGAGTTTACGAACCAAAAAA 698			
Db 701	TGTTGAATTAATGAGTTTACGAACCAAAAAA 739			

LOCUS	BO679163	747 bp	mRNA	linear	EST 15-JUN-2002
DEFINITION	BO679163	AGENCOURT_8034393	NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6213895		
ACCESSION	BO679163				
VERSION	BO679163				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 747)				
AUTHORS	NIH-MGC				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: rgs@nih.gov Tissue Procurement: DCTD/DTF cDNA Library Preparation: Rubin Laboratory DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LCM2380 row: f column: 08 High quality sequence stop: 680.				
FEATURES	Location/Qualifiers				
SOURCE	1..747 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:6213895" /clone_lib="NIH_MGC_112" /tissue_type="melanotic melanoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."				
BASE COUNT	153 a	235 c	227 g	132 t	
ORIGIN					
Query Match	95.7%	Score 676.8;	DB 14;	Length 747;	
Best Local Similarity	99.7%	Pred. No. 3.3e-96;			
Matches 678;	Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;
OY	28	GGGATGGGGCGTGGACGGCCCCCAACTCTTCCTCGGGTCTCCGGGCGGGCTGGCGGACTG	87		
DB	67	GGGATGGGGCGTGGACGGCCCCCAACTCTTCCTCGGGTCTCCGGGCGGGCTGGCGGACTG	126		
OY	88	GGCGTGCAGCGGCTCCCGGCCCACTACAGTTCCTGTTGGACTGGGGGTGCGGACCTG	147		
DB	127	GGCGTGCAGCGGCTCCCGGCCCACTACAGTTCCTGTTGGACTGGGGGTGCGGACCTG	186		
OY	148	GTTGCTCTGACGGAGCGGGGGCCCCCTTCACAGCGACAGCTCCCGGCTCACCCCTGCAC	207		
DB	187	GTTGCTCTGACGGAGCGGGGGCCCCCTTCACAGCGACAGCTCCCGGCTCACCCCTGCAC	246		
OY	208	CGCTGCGGCAATCCCGAATTTTGCCCGCGCGGCCCGACACAGATGACCGCTTGTGTGAC	267		
DB	247	CGCTGCGGCAATCCCGAATTTTGCCCGCGCGGCCCGACACAGATGACCGCTTGTGTGAC	306		
OY	268	ATTCGTGGACGAGGGCCACAGACGGGGGAGAGGCTGTGGAGAGCACTGTGCTGGGCTTT	327		
DB	307	ATTCGTGGACGAGGGCCACAGACGGGGGAGAGGCTGTGGAGAGCACTGTGCTGGGCTTT	366		
OY	328	GGCGCACTGGCAACATCTGCTGGCTGTACTGTGGTGAAGAGAGCGGGCTTGGCTGCAGGA	387		
DB	367	GGCGCACTGGCAACATCTGCTGTTACTGTGTGAAGAGAGCGGGCTTGGCTGCAGGA	426		

QY	388	GATGCGATTTGGCGAATCCGAGACTACGACCCGGCTCCATCGAGACCTATGAGAGGAG	447
DB	427	GATGCGATTTGGCGAATCCGAGACTACGACCCGGCTCCATCGAGACCTATGAGAGGAG	486
QY	448	AAAGCAGCTTCCTCAGCTTACACGGAACGAATTAAGGGGCTTAGTACCCTTCTACACG	507
DB	487	AAAGCAGCTTCCTCAGCTTACACGGAACGAATTAAGGGGCTTAGTACCCTTCTACACG	546
QY	508	GCCCTCATCTCCCTTCCCATGTTTGCATATGGGGCCAGAGATGAAGGGAATGGACTTAA	567
DB	547	GCCCTCATCTCCCTTCCCATGTTTGCATATGGGGCCAGAGATGAAGGGAATGGACTTAA	606
QY	568	GATATTAACCCCTCTAGCTCCCATTTGGCTGGAAGACACTAATATAGCCGCCCTGAGGCA	627
DB	607	GATATTAACCCCTCTAGCTCCCATTTGGCTGGAAGACACTAATATAGCCGCCCTGAGGCA	666
QY	628	GCTCTGATTTGAAGGGAGGCTTGTACTGCTTTGTTGAATTAATGAGTTTACGAAACAA	687
DB	667	GCTCTGATTTGAAGGGAGGCTTGTACTGCTTTGTTGAATTAATGAGTTTACGAAACAA	726
QY	688	AAAAAAAAAAAAAAAAAAAAA 707	
DB	727	AAAAAAAAAAAAAAAAAAAAA 746	
RESULT 5			
LOCUS	B0684685	747 bp	mRNA
DEFINITION	AGNCOCOURT_8061617 NIH_MGC_112 Homo sapiens CDNA clone IMAGE:6214077		
ACCESSION	B0684685		
VERSION	B0684685.1	GI:21797364	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia: Eutheria: Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: gsgabs-remail.nih.gov		
	Tissue Procurement: DCTD/DTP		
	CDNA Library Preparation: Rubin Laboratory		
	CNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA sequencing by: Agencourt Bioscience Corporation		
	clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://Image.lnl.gov		
	Plate: L1CM2380	row: m	column: 22
	High quality sequence spots: 699.		
FEATURES			
SOURCE			
	Location/Qualifiers		
	1..747		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:6214077"		
	/clone_id="NIH_MGC_112"		
	/tissue="melanotic melanoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:		
	EcoRI; CDNA made by oligo-dT priming. Directionally cloned		
	into EcoRI/XhoI sites using the following 5' adaptor:		
	GGGCGAG(G). Library constructed by Ling Hong in the		
	laboratory of Gerald M. Rubin (University of California,		
	Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and		
	Superscript II RT (Life Technologies). Note: this is a		
	NIH_MGC Library."		
BASE COUNT	153 a	235 c	227 g 132 t
ORIGIN			
Query Match	95.7%	Score 676.8;	DB 14; Length 747;
Best Local Similarity	99.7%	Pred. No. 3.3e-96;	

Matches 678; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

28 GCGATGGGGCGTGCAGACCCCACTCTGCTCCGAGGATCCGGGCGCGCTGGCGGACTG 87
 67 GCGATGGGGCGTGCAGACCCCACTCTGCTCCGAGGATCCGGGCGCGCTGGCGGACTG 126
 88 GCGGTCGGGGCGTCCCGCCCACTACAGTTCTGTTGACCTGAGGCTGGGCTGGGACTG 147
 127 GCGGTCGGGGCGTCCCGCCCACTACAGTTCTGTTGACCTGAGGCTGGGCTGGGACTG 186
 148 GTGTCCTGACAG 207
 187 GTGTCCTGACAG 246
 208 GCGGTCGGGGCGTCCCGCCCACTACAGTTCTGTTGACCTGAGGCTGGGCTGGGACTG 267
 247 GCGGTCGGGGCGTCCCGCCCACTACAGTTCTGTTGACCTGAGGCTGGGCTGGGACTG 306
 268 ATGTGAGACAG 327
 307 ATGTGAGACAG 366
 328 GGGGCGACCTGGAC 387
 367 GGGGCGACCTGGAC 426
 388 GATGCCATGCTGTAATTCAGACACACACACACACACACACACACACACACACACAC 447
 427 GATGCCATGCTGTAATTCAGACACACACACACACACACACACACACACACACACAC 486
 448 AAGACAGTCTTCCAGTTCTACACAGCAAAATTAAGGGGCTTGTAGTACCTTCTACAG 507
 487 AAGACAGTCTTCCAGTTCTACACAGCAAAATTAAGGGGCTTGTAGTACCTTCTACAG 546
 508 GGGGCGACCTGGAC 567
 547 GGGGCGACCTGGAC 606
 568 GTATTAACCCCTCTAGCTCCCATTTGGGTGAAGACACAGTGAAGGCTGAGGCTGAGG 627
 607 GTATTAACCCCTCTAGCTCCCATTTGGGTGAAGACACAGTGAAGGCTGAGGCTGAGG 666
 628 GGTCTGATTTGAAGGGGAGGCTTGTACTGCTTGTGTAATTAATTAATTAATTAATTA 687
 667 GGTCTGATTTGAAGGGGAGGCTTGTACTGCTTGTGTAATTAATTAATTAATTAATTA 726
 688 AAAAAAAAAAAAAAAAAAAAAA 707
 727 AAAAAAAAAAAAAAAAAAAAAA 746

RESULT 6
 BM924391 1127 bp mRNA linear EST 12-MAR-2002
 LOCUS 5', mRNA sequence.
 DEFINITION AGENCOURT.6767815 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5760688
 ACCESSION BM924391
 VERSION BM924391.1 GI:19374758
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 1127)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT)
 DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at:
 http://image.llnl.gov
 Plate: ILNLT2808 row: b column: 17
 High quality sequence stop: 565.
 Location/Qualifiers
 1. 1127
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5760688"
 /clone_11b="NIH_MGC_116"
 /lab_host="DH10B"
 /note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH MGC Library."
 BASE COUNT 227 a 362 c 304 g 233 t 1 others
 ORIGIN

Query Match 95.5%; Score 675.4; DB 14; Length 1127;
 Best Local Similarity 99.1%; Pred. No. 4.4e-96;
 Matches 679; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 TGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
 73 TGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 132
 61 GTGCTTCGGGGCGGCTGGGGGAGTGGGCGCTGGGCGGCTGGGCGGCTGGGCGGCTGGG 120
 133 GTGCTTCGGGGCGGCTGGGGGAGTGGGCGCTGGGCGGCTGGGCGGCTGGGCGGCTGGG 192
 121 CTGTTGACCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGGCT 180
 193 CTGTTGACCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGGCT 252
 181 GAGAGTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGG 240
 253 GAGAGTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGG 312
 241 CCCGACGATGACCGCTGCTGACATGCTGACGACGACGACGACGACGACGACGACGAC 300
 313 CCCGACGATGACCGCTGCTGACATGCTGACGACGACGACGACGACGACGACGACGAC 372
 301 GTGGAGTGCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 373 GTGGAGTGCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 432
 361 GTGAGAGAGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 433 GTGAGAGAGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 492
 421 GGTCTCATGAGACCTATGAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 493 GGTCTCATGAGACCTATGAG 552
 481 TAAAGGGCGCTTATGACCTTCTACACAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 553 TAAAGGGCGCTTATGACCTTCTACACAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 612
 541 GCCAGAGATGAAGGAGAGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 600
 613 GCCAGAGATGAAGGAGAGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 672
 601 CACTGAAGTACCCACCCCTGACAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 673 CACTGAAGTACCCACCCCTGACAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732

```

OY 661 GTGATAAATGATTTACGACC 685
DB 733 GTGATAAATGATTTACCAACC 757

RESULT 7
LOCUS BE263806 709 bp mRNA linear EST 13-JUL-2000
DEFINITION 601194169F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3537904 5',
mRNA sequence.
ACCESSION BE263806
VERSION BE263806.1 GI:9137353
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/
1 (bases 1 to 709)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Plate: LICM232 row: j column: 17
High quality sequence stop: 705.
Location/Qualifiers
1..709
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3537904"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site.1: XhoI; Site.2:
EcoRI; cDNA made by oligo-dT priming. Directionality
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT 137 a 233 c 206 g 133 t
ORIGIN
```

```

Query Match 95.0%; Score 671.6; DB 10; Length 709;
Best Local Similarity 98.6%; Pred. No. 2.2e-95;
Matches 688; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
```

```

OY 1 TGACCCGCTGTCCTGTCCTTCCACGATGGCGGACCCCACTTCTCTGG 60
DB 13 TGACCCGCTGTCCTGTCCTTCCACGATGGCGGACCCCACTTCTCTGG 72
OY 61 GTGCTTCGGGCGGCTGCGGAGTGGCGGCTCCCGCCCACTACAGTTC 120
DB 73 GTGCTTCGGGCGGCTGCGGAGTGGCGGCTCCCGCCCACTACAGTTC 132
OY 121 CTGTTGACCTGAGGCTGCGGACCTGTTCTCCCTGACGAGAGCGGCGCCCTCACAAG 180
DB 133 CTGTTGACCTGAGGCTGCGGACCTGTTCTCCCTGACGAGAGCGGCGGCGCCCTCACAAG 192
OY 181 GACAGCTCCCGGCGCTCACCCTGACCGCTGCGCATCCCGACTTCTGCGCGGCGC 240
DB 193 GACAGCTCCCGGCGCTCACCCTGACCGCTGCGCATCCCGACTTCTGCGCGGCGC 252
OY 241 CCCAGCAGATCGACCGCTTGTGAGATCGTGGACGAGGCAAGCGAGGGGAAGGCT 300
DB 253 CCCAGCAGATCGACCGCTTGTGAGATCGTGGACGAGGCAAGCGAGGGGAAGGCT 312
OY 301 GTGGAGAGCACTGTGCTGTGGGCTTTGGCGCACTGGACACACTGTGGCTTTAACCTG 360
DB 313 GTGGAGAGCACTGTGCTGTGGGCTTTGGCGCACTGGACACACTGTGGCTTTAACCTG 372
```

```

OY 361 GTGAAGACCGGGGCTTGGCTGTCAGAGATGCCATTGCTGAATCCGAGCTACGACC 420
DB 373 GTGAAGAGCGGGGCTTGGCTGTCAGAGATGCCATTGCTGAATCCGAGCTACGACC 432
OY 421 GGCTCCATCGAGACCTATGACAGAGAAAGAGCTTCCATTCTACACGAGAAAGAA 480
DB 433 GGCTCCATCGAGACCTATGACAGAGAAAGAGCTTCCATTCTACACGAGAAAGAA 492
OY 481 TAAGGGGCTTGTATACCTTCTACAGGCGCTCACTCCCTTCCCATTTGTGATGG 540
DB 493 TAAGGGGCTTGTATACCTTCTACAGGCGCTCACTCCCTTCCCATTTGTGATGG 552
OY 541 GCCAGATGAAAGGAGTGAAGTAAAGTATTAACCCCTTACGTCCTCCATTGGCTGAAGA 600
DB 553 GCCAGATGAAAGGAGTGAAGTAAAGTATTAACCCCTTACGTCCTCCATTGGCTGAAGA 612
OY 601 CACTGAAGTACGCCACCCCTGACGAGAGTCTCTGATTTGAAGGGAGGCTTACGCTT 660
DB 613 CACTGAAGTACGCCACCCCTGACGAGAGTCTCTGATTTGAAGGGAGGCTTACGCTT 671
OY 661 GTGATAAATGATTTACGAAACCAAAAAA 698
DB 672 GTGATAAATGATTTACGAAACCAAAAAA 709
```

```

RESULT 8
LOCUS B1598433 725 bp mRNA linear EST 07-SEP-2001
DEFINITION 603246943F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5294575 5',
mRNA sequence.
ACCESSION B1598433
VERSION B1598433.1 GI:15491372
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
```

```

REFERENCE NIH-MGC http://mgc.ncl.nih.gov/
1 (bases 1 to 725)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
```

```

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: ILAM11745 row: e column: 12
High quality sequence stop: 668.
Location/Qualifiers
1..725
```

```

FEATURES
SOURCE
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5294575"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROP 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
```

```

BASE COUNT 145 a 233 c 216 g 131 t
ORIGIN
```

Query Match 91.3%; Score 645.8; DB 13; Length 725;
 Best Local Similarity 98.7%; Pred. No. 2.2e-91;
 Matches 672; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

25 CCAGCGATGGGCTGCGAGCCGCCCACTTCTCTGGTGTCTCCGGCCGGCTGGCGGA 84
 |||||
 34 CCAGCGATGGGCTGCGAGCCGCCCACTTCTCTGGTGTCTCCGGCCGGCTGGCGGA 93
 |||||
 85 CTGGCGCTGGCGGCTGCGAGCCGCCCACTTCTCTGGTGTCTCCGGCCGGCTGGCGCA 144
 |||||
 94 CTGGCGCTGGCGGCTGCGAGCCGCCCACTTCTCTGGTGTCTCCGGCCGGCTGGCGCA 153
 |||||
 145 CTGGTGTCTCTGAGCGAGCGCGGGCCCTCAGAGCGAGCTGCCCCGCTCACCCTG 204
 |||||
 154 CTGGTGTCTCTGAGCGAGCGCGGGCCCTCAGAGCGAGCTGCCCCGCTCACCCTG 213
 |||||
 205 CACCGCTGGGCTGCGAGCGCGGGCCCTCAGAGCGAGCTGCCCCGCTCACCCTG 264
 |||||
 214 CACCGCTGGGCTGCGAGCGCGGGCCCTCAGAGCGAGCTGCCCCGCTCACCCTG 273
 |||||
 265 CAGATGCTGGGCTGCGAGCGCGGGCCCTCAGAGCGAGCTGCCCCGCTCACCCTG 324
 |||||
 274 CAGATGCTGGGCTGCGAGCGCGGGCCCTCAGAGCGAGCTGCCCCGCTCACCCTG 333
 |||||
 325 TTTGGCGGCTGCGAGCGCGGGCCCTCAGAGCGAGCTGCCCCGCTCACCCTG 384
 |||||
 334 TTTGGCGGCTGCGAGCGCGGGCCCTCAGAGCGAGCTGCCCCGCTCACCCTG 393
 |||||
 385 GAGAGTGGCTGCTGTAATCCAGAGCTGCGAGCGCGGGCTGCGAGCTGCGAG 444
 |||||
 394 GAGAGTGGCTGCTGTAATCCAGAGCTGCGAGCGCGGGCTGCGAGCTGCGAG 453
 |||||
 445 GAGAGTGGCTGCTGTAATCCAGAGCTGCGAGCGCGGGCTGCGAGCTGCGAG 504
 |||||
 454 GAGAGTGGCTGCTGTAATCCAGAGCTGCGAGCGCGGGCTGCGAGCTGCGAG 513
 |||||
 505 CAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 564
 |||||
 514 CAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 573
 |||||
 565 AAGATTTAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 624
 |||||
 574 AAGATTTAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 633
 |||||
 625 GAGAGTGGCTGCTGTAATCCAGAGCTGCGAGCGCGGGCTGCGAGCTGCGAG 684
 |||||
 634 GAGAGTGGCTGCTGTAATCCAGAGCTGCGAGCGCGGGCTGCGAGCTGCGAG 691
 |||||
 685 CAAAAAATTTAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705
 |||||
 692 CCACGACAAAAAATTTAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 712
 |||||

RESULT 9
 B0892874 904 bp mRNA linear EST 16-AUG-2002

LOCUS B0892874 5' mRNA sequence.

DEFINITION B0892874.1 GI:22284888

ACCESSION B0892874

VERSION B0892874.1

KEYWORDS human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Creniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 904)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@pds-remail.nih.gov
 Tissue Procurement: DCID/DPF

CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LCM231 row: 3 column: 06
 High quality sequence stop: 638.

FEATURES
 source
 1. 904
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6090341"
 /clone_lib="NIH_MGC_112"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pORF7; Site: 1; XhoI; Site: 2;
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 SuperScript II RT (Life Technologies). Note: this is a
 NIH-MGC Library."

BASE COUNT 185 a 270 c 279 g 169 t 1 others
 ORIGIN

Query Match 87.9%; Score 621.4; DB 14; Length 904;
 Best Local Similarity 96.2%; Pred. No. 1.2e-87;
 Matches 658; Conservative 0; Mismatches 22; Indels 4; Gaps 2;

28 GCGATGGGCTGAGCGCCCACTTCTCTGGTGTCTCCGGCCGGCTGGCGGACTG 87
 |||||
 67 GCGATGGGCTGAGCGCCCACTTCTCTGGTGTCTCCGGCCGGCTGGCGGACTG 126
 |||||
 88 GCGCTGCGGCTGCGAGCGCGGGCCCTCAGAGCGAGCTGCCCCGCTCACCCTG 147
 |||||
 127 GCGCTGCGGCTGCGAGCGCGGGCCCTCAGAGCGAGCTGCCCCGCTCACCCTG 186
 |||||
 148 GGTGCTGCGAGCGAGCGCGGGCCCTCAGAGCGAGCTGCCCCGCTCACCCTG 207
 |||||
 187 GGTGCTGCGAGCGAGCGCGGGCCCTCAGAGCGAGCTGCCCCGCTCACCCTG 246
 |||||
 208 GCGCTGCGGCTGCGAGCGCGGGCCCTCAGAGCGAGCTGCCCCGCTCACCCTG 267
 |||||
 247 GCGCTGCGGCTGCGAGCGCGGGCCCTCAGAGCGAGCTGCCCCGCTCACCCTG 306
 |||||
 268 ATGCTGAGAGCGAGCGAGCGGGGAGAGCTGTGGAGTGAAGTGTGCTGCTGCTGCTG 327
 |||||
 307 ATGCTGAGAGCGAGCGAGCGGGGAGAGCTGTGGAGTGAAGTGTGCTGCTGCTGCTG 366
 |||||
 328 GCGCGACTGCGAGCGAGCGGGGAGAGCTGTGGAGTGAAGTGTGCTGCTGCTGCTG 387
 |||||
 367 GCGCGACTGCGAGCGAGCGGGGAGAGCTGTGGAGTGAAGTGTGCTGCTGCTGCTG 426
 |||||
 388 GATGCGATGCGAGCGAGCGGGGAGAGCTGTGGAGTGAAGTGTGCTGCTGCTGCTG 447
 |||||
 427 GATGCGATGCGAGCGAGCGGGGAGAGCTGTGGAGTGAAGTGTGCTGCTGCTGCTG 486
 |||||
 448 AAGAGCTTCTCAGTTCTACAGCGAGCAAAATTAAGGGGCTTTATACCTCTCTACAG 507
 |||||
 487 AAGAGCTTCTCAGTTCTACAGCGAGCAAAATTAAGGGGCTTTATACCTCTCTACAG 546
 |||||
 508 GCGCTGCTGCGAGCGAGCGGGGAGAGCTGTGGAGTGAAGTGAAGTGAAGTGAAGTGA 567
 |||||
 547 GCGCTGCTGCGAGCGAGCGGGGAGAGCTGTGGAGTGAAGTGAAGTGAAGTGAAGTGA 606
 |||||
 568 GATTTAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 627
 |||||
 607 GATTTAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666
 |||||
 628 GGTCC--TGATGAAGGGGAGGAGCTGTACTG--CTTTGTTGAATTAATGATTTACGAA 683
 |||||

Db 667 AGGCCCTGATTGTAAGGGAGGCTTGACTGCTTTGTTGGAATATGACTTTTACGA 726

QY 684 CCAAAAAAAAAAAAAAAAAAAAAA 707

Db 727 AACCAAGCAAAAAAAAAAAGAA 750

RESULT 10

BF970174 712 bp mRNA linear EST 22-JAN-2001

LOCUS 602273385F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4361564 5', mRNA sequence.

ACCESSION BF970174

VERSION BF970174.1 GI:12337389

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 712)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Plate: LLM10004 row: 1 column: 21

High quality sequence stop: 657.

Location/Qualifiers

1. 712

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4361564"

/clone_lib="NIH_MGC_84"

/tissue_type="adrenal cortex carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 135 a 231 c 216 g 130 t

ORIGIN

Query Match 86.1%; Score 608.6; DB 12; Length 712; Best Local Similarity 94.1%; Pred. No. 1.3e-85; Matches 643; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 25 CCACGATGGGCGCTGACAGCCGCCCACTTCTCCGCGGCTCCGCGCGGCGGCGGCA 84

Db 23 CCACGATGGGCGCTGACAGCCGCCCACTTCTCCGCGGCTCCGCGCGGCGGCGGCA 82

QY 85 CTGGCGCTGGCGCGGCTCCGCCCACTTCTCCGCGGCTCCGCGGCTCCGCGGCGGCA 144

Db 83 CTGGCGCTGGCGCGGCTCCGCCCACTTCTCCGCGGCTCCGCGGCTCCGCGGCGGCA 142

QY 145 CTGGTGTCCCTGACGAGCGCGGCGGCGGCTCCACAGCAGACGCTCCCGGCTCACCTG 204

Db 143 CTGGTGTCCCTGACGAGCGCGGCGGCGGCTCCACAGCAGACGCTCCCGGCTCACCTG 202

QY 205 CACGCGCTGGCGATCCCGGACTTCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 264

Db 203 CACGCGCTGGCGATCCCGGACTTCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 262

QY 265 CAGATCGTGAAGGCGCAACGCAAGCGGCGGAGAGGCTGTGGAGTGCACCTGTGCTGGCG 324

Db 263 CAGATCGTGAAGGCGCAACGCAAGCGGCGGAGAGGCTGTGGAGTGCACCTGTGCTGGCG 322

QY 325 TTGGCCGCACTGGCACCACATGCTGGCTTACCTGCTGTAAGGAGCGGGGCTTGCTGCA 384

Db 323 TTGGCCGCACTGGCACCACATGCTGGCTTACCTGCTGTAAGGAGCGGGGCTTGCTGCA 382

QY 385 GGAGATGCCATTGCTGAAATCCGAGACTAGCAGCCGGCTCCATGAGACCTATGACAG 444

Db 383 GGAGATGCCATTGCTGAAATCCGAGACTAGCAGCCGGCTCCATGAGACCTATGACAG 442

QY 445 GAGAAAGCAGCTTTCACCTTACAGGCAAGAAATAGGGGCTTGTAGTACCTTAC 504

Db 443 GAGAAAGCAGCTTTCACCTTACAGGCAAGAAATAGGGGCTTGTAGTACCTTAC 502

QY 505 CAGGCCCTCACTCCCTTCCCATGTTGCGATGGGGCCAGAGATGAAGGAAAGTGA 564

Db 503 CAGGCCCTCACTCCCTTCCCATGTTGCGATGGGGCCAGAGATGAAGGAAAGTGA 562

QY 565 AAAGTATTAACCTCTGATGCTCCATGTCGTAAGACACTGAATGACCCCTGAC 624

Db 563 AAAGTATTAACCTCTGATGCTCCATGTCGTAAGACACTGAATGACCCCTGAC 621

QY 625 GCAGTCTGATTAAGGAGGAGCTTGACTGCTTTGTAATAATGATTTACGAC 684

Db 622 GCAGTCTGATTAAGGAGGAGCTTGACTGCTTTGTAATAATGATTTACGAC 681

QY 685 CAAAAAAAAAAAAAAAAAAAAA 707

Db 682 CACGAACCAAGCAACCAACCA 704

RESULT 11

BG831257 828 bp mRNA linear EST 22-MAY-2001

LOCUS 602766474F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4908258 5', mRNA sequence.

ACCESSION BG831257

VERSION BG831257.1 GI:14178844

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 828)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Plate: LLM1814 row: d column: 19

High quality sequence stop: 610.

Location/Qualifiers

1. 828

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4908258"

/clone_lib="NIH_MGC_42"

/tissue_type="epithelioid carcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

"

Query Match	83.88;	Score 592.6;	DB 12;	Length 828;
Best Local Similarity	99.08;	Pred. No. 3.8e-83;		
Matches 617;	Conservative 4;	Mismatches 2;	Gaps 2;	

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM1732 row: e column: 12
High quality sequence stop: 638.

FEATURES	Location/Qualifiers
source	1. .639

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:486035"
/clone_1lb="NH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCAACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

```

Query Match	83.1%	Score 587.4	DB 12	Length 639
Best Local Similarity	99.8%	Pred. 2.8e-82		
Matches 588	Conservative	0	Mismatches 1	Indels 0
				Gaps 0

Db	313	CTTTGGGCGCACTGGCACACATGTGGGCTGTACTGTTAAGGAGGAGGCGGGCTGGCTGC	372
OY	384	AGAGATAGCCATTCTCTGAATTCGAGACATACGACCCGGCTGCATGACACTATAGCA	443
Db	373	AGGAGATGCCATTCTCTGAATTCGAGACATACGACCCGCTGCATGACACTATAGCA	432
OY	444	GGAGAAAGCACTCTTCAGTTCTACACGCAACGAAATAAGGGGCTTAACTACCTTCTA	503
Db	433	GGAGAAAGCACTCTTCAGTTCTACACGCAACGAAATAAGGGGCTTAACTACCTTCTA	492
OY	504	CCAGGCCCTACATCCCTTCCCATCTTGTTCGATGGGGCCAGAGATGAAGGAATGGAC	563
Db	493	CCAGGCCCTACATCCCTTCCCATCTTGTTCGATGGGGCCAGAGATGAAGGAATGGAC	552
OY	564	TAAAGTATTAACCCCTCTAGCTCCATCTGGCTGTAAGACATCGAAGTAGCCACCTGTCA	623
Db	553	TAAAGTATTAACCCCTCTAGCTCCATCTGGCTGTAAGACATCGAAGTAGCCACCTGTCA	612
OY	624	GGC-AGGTCTGATTGAAGGGA	645
Db	613	GGCAAGGTCTGATCAGCGGCGA	635

RESULT 12	LOCUS	DEFINITION
639 bp	60277654.F1	60277654.F1 NIH_MGC_15 Homo sapiens
linear		CDNA clone IMAGE:4806035 5',
EST 22-MAY-2001		mRNA sequence.

VERSION	BG823185.1	GI:14170772
KEYWORDS	EST.	
SOURCE	human.	

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 639)

AUTHORS NIH-MGC <http://mgc.>

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
Unpublished (1999)
COMMENT

CONTACT: ROBERT STRAUSBERG, Ph.D.
Email: roberts@ucla.edu

Email: cygabs-remail.nln.gov
 Tel: +31 20 674 4200 ext: 3400

Issue Procurement: AIC
CDNA Library Preparation: 1 and 2000/2010

COMM LIBRARY REFERENCE: Ling Hong/Kuolin Laboratory

RESULT 13
AI681916

QY 481 TAAAGGGGCTTAGTACCTCTTCTACAGAGCCCTCACTCCCTCCCAATGTTGATGG 540

Db 531 TAAAGGGGCTTAGTACCTCTTCTACAGAGCCCTCACTCCCTCCCAATGTTGATGG 590

QY 541 GCCAGAGATGAAGGAAGTGACTAAAGTATTAAACCTCTAGCTCCCA 589

Db 591 GCCAGAGATGAAGGAAGTGACTAAAGTATTAAACCTCTAGCTCCCA 639

RESULT 13
AI681916

LOCUS	602 bp	RNA	linear	EST 16-DEC-1999
LOCUS	AI681916			
DEFINITION	tx50f02.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:2273019 3'			
ACCESSION	AI681916			
VERSION	AI681916.1	GI:4892098		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.			
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index			
JOURNAL	Unpublished (1997)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILMIL at: www.bio.lml.gov/db/rrp/image/image.html Insert length: 681 Std Error: 0.00 Seq primer: -40bp from Glibco High quality sequence stop: 495. Location/Qualifiers 1. 602 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2273019" /clone_lib="NCI-CGAP_Lu24" /tissue_type="carcinoid" /lab_host="DH10B" /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI-CGAP_Lu24 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldi."			
FEATURES				
source				
BASE COUNT	109 a	202 c	179 g	110 t
ORIGIN				
Query Match	82.6%;	Score 583.8;	DB 9;	length 602;
Best Local Similarity	99.3%;	Pred. No. 1e-01;		
Matches	596;	Conservative	0;	Mismatches 3; Indels 1; Gaps 1;
OY	42	GGCCCCCACTTCTCTGGGTGCTTCGCGGCGCGGTGGCGGGACTGGCGCGGGCT	101	
Db	3	GGCGCGCAACTTCTCTGGGTGCTTCGCGGCGCGGTGGCGGGACTGGCGCGGGCT	62	
OY	102	CCCCGCCACATACAGATTCTCTGTTGGACCTGGGGCTGGCGACACTGTGTCTCCAGCGA	161	
Db	63	CCCCGCCACATACAGATTCTCTGTTGGACCTGGGGCTGGCGACACTGTGTCTCCAGCGA	122	
OY	162	GGGGGGGGGGGGCTGCACAGACAGGTGGCCGGCGGCTGCACCCCTGCACCGCGCCGATGCC	221	
Db	123	GGGGGGGGGGGGCTGCACAGACAGGTGGCCGGCGGCTGCACCCCTGCACCGCGCGATGCC	182	
OY	222	CGACTTTCGCCCGCGCGGGCCCCGACCAATGCACGCTTCGTGAGATCGTGGAGAGAGC	281	
Db	183	CGACTTTCGCCCGCGCGGGCCCCGACCAATGCACGCTTCGTGAGATCGTGGAGAGAGC	242	
OY	282	CAACGACGAGGAGAGGAGCTGTGGAGATGCATGTGCTCTGGGCTTTGGCGCAGCTGAC	341	
Db	243	CAACGACGAGGAGAGGAGCTGTGGAGATGCATGTGCTCTGGGCTTTGGCGCAGCTGAC	302	
OY	342	CATGCTGGCCGTTCACCTGGTGCA-GGAGCGGGGCTTGGCTGCAGAGATGCATTGCTG	400	

DB	Accession	Version	KeyWords	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	303	CATGCTGGCCCTGTACCTGTGTGAANGAGCGGANGCTTGAGCTGCAGAGATGTCATTGGCG								
QY	401	AAATCCGAGACTACGACCCGGGCTCCATTCGABACTATGAGAGAGAAAGACGCTTCC								
Db	363	AAATCCGAGACTACGACCCGGGCTCCATTCGABACTATGAGAGAGAAAGACGCTTCC								
QY	461	AGTTCTACAGGAGCAAAATGAAGGGGCTTGTATACCTTTCACAGAGCCCTCACCTCCC								
Db	423	AGTTCTACAGGAGCAAAATGAAGGGGCTTGTATACCTTTCACAGAGCCCTCACCTCCC								
QY	521	TTTCCCATGTGTGTGCATGGGGCCAGAGATGAAGGAGAGTGACATAAGTATTAACCTTC								
Db	483	TTTCCCATGTGTGTGCATGGGGCCAGAGATGAAGGAGAGTGACATAAGTATTAACCTTC								
QY	581	TAGCTCCATTGGCTGGAAGACACTAATAGCCACCCCTGCAGGACAGCTCTGATTGAA								
Db	543	TAGCTCCATTGGCTGGAAGACACTAATAGCCACCCCTGCAGGACAGCTCTGATTGAA								
RESULT 14										
LOCUS	BO214823			931 bp	mRNA	linear	EST 02-MAY-2002			
DEFINITION	AGNCOURT_759286 NIH_MGC_72 Homo sapiens				cdna	clone	IMAGE:6063586			
ACCESSION	BO214823									
VERSION	BO214823.1			GI:20396223						
KEYWORDS	EST.									
SOURCE	human.									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.									
TITLE	NIH-MGC http://mgc.ncl.nih.gov/.									
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)									
COMMENT	Unpublished (1999)									
	Contact: Robert Strausberg, Ph.D.									
	Email: cgrabbs-remail.nih.gov									
	Tissue Procurement: ATCC/DCTP/DTF									
	CDNA Library Preparation: Life Technologies, Inc.									
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)									
	DNA sequencing by: Agencourt Bioscience Corporation									
	Clone distribution: MGC clone distribution information can be									
	found through the I.M.A.G.E. Consortium/LNLN at:									
	http://Image.lnl.gov									
	Plate: LLM1336			row: 0	column: 11					
	High quality sequence stop: 563.									
FEATURES										
source	1..931									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
	/clone="IMAGE:6063586"									
	/clone_1id="NIH.MGC.72"									
	/lab_type="melanotic melanoma"									
	/lab_host="DH10B (phage-resistant)"									
	/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;									
	Site_2: SalI; Cloned unidirectionally. Primer: oligo dt.									
	Average insert size 2 kb. Library constructed by Life									
	Technologies."									
BASE COUNT	198 a	284 c	270 g	179 t						
ORIGIN										
Query Match	78.6%;	Score 555.6;	DB 14;	Length 931;						
Best Local Similarity	99.3%;	Pred. No. 2e-77;								
Matches 558;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;</						

251 TCACCGCTTCGTGACATGCTGACAGAGCCAAACGACGCGGGAGAGCTGTGGAGTGC 310
 121 TCACCGCTTCGTGACATGCTGACAGAGCCAAACGACGCGGGAGAGCTGTGGAGTGC 180
 311 ACCTGCTCTGGGCTTTGGCCGACACTGCGACACTGCTGGCTTACCTGTGAGAGGC 370
 181 ACTGCTCTGGGCTTTGGCCGACACTGCGACACTGCTGGCTTACCTGTGAGAGGC 240
 371 GGGGCTTGGCTGACAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 430
 241 GGGGCTTGGCTGACAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 300
 431 AGACCTATGACAGAGAGAAAGAGTCTTCCAGTTTACACGACGAAATTAAGGGGCT 490
 301 AGACCTATGACAGAGAGAAAGAGTCTTCCAGTTTACACGACGAAATTAAGGGGCT 360
 491 TAGTACCTTCTACAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 550
 361 TAGTACCTTCTACAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 551 AAGGAGAGTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 610
 421 AAGGAGAGTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 480
 611 GCCACCCCTGACAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 670
 481 GCCACCCCTGACAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 671 TAGTATTCAGAACCAAAAAA 692
 541 TAGTATTCAGAACCAAAAAA 562

RESULT 15
 LOCUS B1761318 584 bp mRNA EST 25-SEP-2001
 DEFINITION B1761318 NIH-MGC_116 Homo sapiens cDNA clone IMAGE:5184523 5',
 mRNA sequence.
 ACCESSION B1761318
 VERSION B1761318.1 GI:15752896
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NIH-MGC http://mhc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM11460 row: k column: 20
 High quality sequence stop: 583.
 Location/Qualifiers
 1..584
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5184523"
 /clone_11b="NIH-MGC_116"
 /lab_host="DH10B"
 /note="Organ: pooled colon, kidney, stomach; Vector:
 pCMV-SPORT6; Site: 1; NotI; Site: 2; EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 26 yo male, 49 yo
 female, 71 yo male colon; 46 yo male kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library 18

oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb.
 Insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH-MGC Library."

BASE COUNT 90 a 209 c 181 g 104 t
 ORIGIN

Query Match 77.9%; Score 551; DB 13; Length 584;
 Best Local Similarity 100.0%; Pred. No. 1.3e-76;
 Matches 551; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TGACCGCTTCGTGACATGCTGACAGAGCCAAACGACGCGGGAGAGCTGTGGAGTGC 60
 33 TCACCGCTTCGTGACATGCTGACAGAGCCAAACGACGCGGGAGAGCTGTGGAGTGC 92
 61 GTGCTTCGGGCGGGCTGGGAGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 120
 93 GTGCTTCGGGCGGGCTGGGAGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 152
 121 CTGTTGAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 180
 153 CTGTTGAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 212
 181 GACAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 240
 213 GACAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 272
 241 CCCGACAGATGACAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 273 CCCGACAGATGACAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 332
 301 GTGAGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 333 GTGAGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 392
 361 GTGAGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 393 GTGAGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 452
 421 GGCTCATGAGACCTATGAG 480
 453 GGCTCATGAGACCTATGAG 512
 481 TAGGGGCTTACCTTCTACAGAGGCTTACCTTCTACAGAGGCTTACCTTCTACAGAGGCTT 540
 513 TAGGGGCTTACCTTCTACAGAGGCTTACCTTCTACAGAGGCTTACCTTCTACAGAGGCTT 572
 541 GCCAGAGATGA 551
 573 GCCAGAGATGA 583

Search completed: November 4, 2002, 09:24:20
 Job time : 2075 secs

This Page Blank (uspto)